New Insights Into the Nature of the HIV-2 Reservoir

NEW INSIGHTS INTO THE NATURE OF THE HIV-2 RESERVOIR

A dissertation presented

by

Fernanda de Araújo Ferreira

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Virology

Harvard University

Cambridge, Massachusetts

September 2019

Dissertation Advisor: Daniel R. Kuritzkes Author: Fernanda de Araújo Ferreira

New Insights into the Nature of the HIV-2 Reservoir

Abstract

Acquired immunodeficiency syndrome (AIDS) can be caused by two lentiviruses: HIV-1, which is responsible for the global AIDS pandemic, and HIV-2, which is endemic in west Africa. In Chapter 1, we show how the two viruses share a similar genetic structure and replication cycle, but are markedly different in their clinical outcomes.

In those infected with HIV-1 and HIV-2, the latent reservoir is a major roadblock in the search for a cure. While the nature of the HIV-1 reservoir has and continues to be investigated, the HIV-2 reservoir has only been recently probed. In Chapter 2, we expand on previous work and characterize the proviral landscape of 9 HIV-2-infected study participants. We find that, like HIV-1, the HIV-2 integrated provirus is mostly defective, with large deletions being the most common defect.

One potential HIV-1 cure strategy proposed is "shock and kill," which involves the use of latency reversal agents (LRAs) to reactivate the HIV provirus and eliminating HIV-infected cells via virus-mediated cytotoxicity or immune clearance. In Chapter 3, we explore the possibilities of "shock and kill" for individuals latently-infected with HIV-2. We performed *ex vivo* reactivation experiments on PBMCs isolated from HIV-2 study participants and find that, when we control for DNA input, HIV-1 and HIV-2 PBMCs produce comparable amounts of cell-associated RNA. However, HIV-1 PBMCs produce higher levels of supernatant RNA compared to HIV-2 PBMCs for the majority of the LRA conditions.

In Chapter 4, we discuss the need for a consensus proviral landscape characterization pipeline. By exploring the proviruses of 13 HIV-1-infected study participants, we find that the majority of proviruses harbor two or more types of defects, which is not reflected in current pipelines. This discrepancy leads to differences in both individual and cumulative proviral landscapes.

Taken together, our studies highlight that the nature of the proviral landscape is determined more by host factors than viral factors, and that the differences in virions production post-reactivation of HIV-1- and HIV-2-infected PBMCs is not due to differences in DNA or RNA abundance, but rather to a previously undescribed post-transcriptional block.

Acknowledgements

HIV-2 forced me to be a better assay developer, taught me about the perverse pleasure that can be derived from tinkering with an assay till you reach that sweet R^2 of 0.99, and forced me to be more open about failure. These are all lessons that are valuable not just in lab, but also in life (well, maybe not the assay development part). Holding vials of blood from HIV-2-infected study participants gave me an unfaltering respect to the study participants that have given me their time and literal blood. You may be the "who?" of the HIV world, but you are a "them" in my heart.

I also want to thank Dan Kuritzkes and Athe Tsibris, without which I would never have ended up working with HIV-2. Not only did you place the best virus in my hands and let me take complete possession of this project, you both also created a lab environment to which I was happy to return to day-after-day even when a project had stalled or was falling apart. PIs are the ones who set the tone for the lab, who determine what environment we do our research in, and our lab has always been open, inviting, a place where friendships were made and with an endearing level of kookiness that made it a fun place to be, and for that I cannot thank you both enough.

Athe, there are many ways to describe you. Exceedingly tall. Obnoxiously enthusiastic about science. Orator of phrases like "life is about expectation management" when I said the qPCR assay would be up and running in a month (lol). But beyond that, every graduate student deserves a PI who will look out for them, both as a scientist and a human, and you've done that for me. Which means the best way to describe you is as a friend.

The K lab has always had the best technicians. To the first wave of techs–Hayat, Sara, Emily, Lisa, Amanda and Allison–thank you for inviting me into the lab with open arms and for indulging my love of group Halloween costumes. To the second wave of techs–Stephanie, Olivia and Qianjing–thank you for being such ridiculous paragons of organization and productivity, and

for hiding 100 pictures of Jeff Goldblum throughout the lab. To the third wave of techs–Heather, Kendyll, Jesse and Colline–thank you for keeping me up-to-date on what the youths are doing/saying/watching and, in doing so, making me aware of my age.

To the members of my DAC–Todd Allen, Marty Hirsch and Joe Sodroski–in your honesty and support, you gave me the push I needed to start up a new project. I went from dreading DAC meetings, to looking forward to them and all the incredible feedback I knew I would receive.

Radwa and Phil, my fellow travelers on the PhD journey. I could not choose a better pair to cry in front of and vent about the current state of science with. To my fellow G6s and the Virology Program: I cannot tell you how nice it has been to have spent the last few years surrounded by people who also have a favorite virus!

I owe my love of viruses to Renato Oliveira Resende, Tatsuya Nagata and Fernando Lucas Melo of the Universidade de Brasilia (UnB), who taught me all I know about bench science and instilled in me an appreciation for *Tomato spotted wilt virus* that continues till this day.

I want to thank the Brazilian National Council for Scientific and Technological Development (CNPq) for three years of scholarships and the Lemann Foundation for additional financial support. Things are pretty bad now, but Brazilian science has always muddled through its low moments and triumphed despite them.

To my family and friends, who always took my calls, and told me a good cry and a shower could cure anything (you were right!). To my sister, Carolina, who called at midnight just so I could hear the crowds in Toronto lose their minds over the Raptors, and will always bring me up to speed about what's happening in the MCU.

And finally, to my parents, who tricked me into thinking that moving to a foreign country to pursue a PhD would be easy.

Table of Contents

List of Figures

CHAPTER 1: INTRODUCTION

HIV-2: Its Replication Cycle and Reservoir

Introduction

> *Previous studies of HTLV-III antibodies in Africa have focused on central Africa, where AIDS has been observed. Our study suggests that people in Senegal, where AIDS has not yet been reported, have also been exposed to viruses of the HTLV-III class. However, the virus we found in some healthy prostitutes and surgical inpatients is more closely related to STLV-III$_{AGM}$ than to reference strains of HTLV-III.*
>
> - Barin *et al.*,1985 (1)

When Barin and colleagues looked for signs of HTLV-III[1] infection in Senegal in western Africa, they found something new. The sera of 20 of the 289 sex workers and 5 of the 122 surgical patients they surveyed were positive for antibodies against HTLV-III by ELISA, showing for the first-time evidence of an HTLV-III-like virus in Senegal. When further analyzed via western blot, these participants' sera were also reactive to the p32 and p24 proteins of Simian T-Lymphotropic Retrovirus III from African green monkeys (STLV-III$_{AGM}$, now SIV$_{agm}$). These blots were strikingly different from what they observed with sera of infected individuals from Burundi, which is in central Africa, and the USA: no reaction to p32 or p24 of STLV-III$_{AGM}$. For Barin *et al.*, this finding suggested the presence of a new virus, one that was related to both STLV-III and HTLV-III and, most interestingly, did not appear to cause AIDS.

This last point was key. A few months later, when this new virus, which would eventually be called HIV type 2 (HIV-2), was isolated from the same group of STLV-III$_{AGM}$ antibody-positive

[1] HTLV-III was one of the names used for the virus we now call Human Immunodeficiency Virus type 1 (HIV-1). Lymphadenopathy-Associated Virus (LAV) was another early name, given by the French group that isolated it. In 1986, the International Committee on the Taxonomy of Viruses decreed that HIV-1 would be the official name for the virus.

individuals from Senegal, the authors would reiterate that the absence of AIDS in HIV-2-infected individuals imbued this virus with the potential to both better understand HIV-1 and aid in the development of an AIDS vaccine (2). Even now, when it is clear that in some individuals, infection with HIV-2 does lead to AIDS, this virus continues to be primarily studied as a means for furthering our understanding of HIV-1 pathogenicity (3,4).

In the 33 years since HIV-2 was isolated, studies have revealed a virus that can best be described as "similar but different" to HIV-1 in terms of its replication cycle, yet substantially different in terms of its epidemiology (a broad strokes comparison of the two human retroviruses can be found in Table 1). The major difference between the two viruses and the aspect of HIV-2 infection that has kept us studying it is its lower pathogenicity, which sees the majority of those infected with the virus behaving as long term non-progressors (LTNPs). This feature brings with it the continued hope that investigating HIV-2 will lead the field to a better understanding of the host and viral factors that contribute to its comparably lower pathogenicity.

In this chapter, we will summarize what we know about HIV-2 replication cycle and epidemiology, highlighting the ways in which it differs to HIV-1 and mentioning areas that require further investigation or confirmatory experiments. Finally, we will discuss what is currently known about the HIV-2 reservoir, which is the focus of this dissertation.

*

Barin *et al*. highlighted the apparent healthiness of the study participants they surveyed, but months later it became clear that like HIV-1, HIV-2 could also cause AIDS (1,3,4). Despite both viruses being the etiologic agents behind AIDS and having a similar genetic structure–two long terminal repeats, three major genes (gag, pol, env) and a cadre of genes encoding additional proteins–they only share about 40% overall nucleotide sequence homology (Figure 1) (5). HIV-2

Table 1. HIV-1 vs HIV-2: A Brief Comparison of HIV-1 and HIV-2

	HIV-1	HIV-2
Epidemiology		
Clinical Illness		
Acute Phase	High viral loads and decreases in the number of $CD4^+$ T cells	No data, but assumed to be similar to HIV-1
Asymptomatic Stage	Usually lasts around 10 years	Lasts over 10 years
AIDS	Low CD4 counts and opportunistic infections	Similar to HIV-1
Viral Load	In ART naïve participants: 10^4 to 10^6 RNA copies per mL	In ART naïve study participants: 30 times lower
Distribution	Global	West Africa and countries with links to West Africa, especially Portugal and France
Transmission	Through bodily fluids: sexual, blood borne and MTCT	Same routes as HIV-1, but less efficient
Replication Cycle		
Genome	*Gag*, *pol* and *env* are the structural proteins, but there are also the regulatory proteins and accessory proteins (*vif*, *vpr*, *vpu* and *nef*)	Same structural, regulatory and accessory proteins, except for *vpu*. In its place, HIV-2 carries *vpx*.
Viral Particle	Cone-shaped capsid containing two linked genomic RNAs; surrounded by an envelope	Similar to HIV-1 and indistinguishable through microscopy
Entry Receptors & Coreceptors	CD4 is the main entry receptor, with CCR5 and CXCR4 as the primary co-receptors	Similar to HIV-1, though there are more cases of CD4-independent entry *in vitro* and HIV-2 shows higher promiscuity for alternative co-receptors (also *in vitro*)
Reverse Transcription	HIV-1 RT appears to have a higher rate of dNTP incorporation	HIV-2 carries *vpx*, which counteracts SAMHD1 allowing it to reverse transcribe in cells with low dNTP content
Integration	Preferentially integrates into transcriptional units	Similar to HIV-1, though HIV-2 appears to favor integration in the reverse direction
Transcription: Enhancer Elements	Typically, two NF-kB, three SP-1, two AP-1 and one Ets binding sites	Typically, three SP-1 and only 1 functional NF-kB binding sites
Packaging	Involves *cis*- and *trans*-acting elements	Proposed packaging mechanism is primarily *cis*-acting, with Gag proteins being translated from RNA genomes that they then package

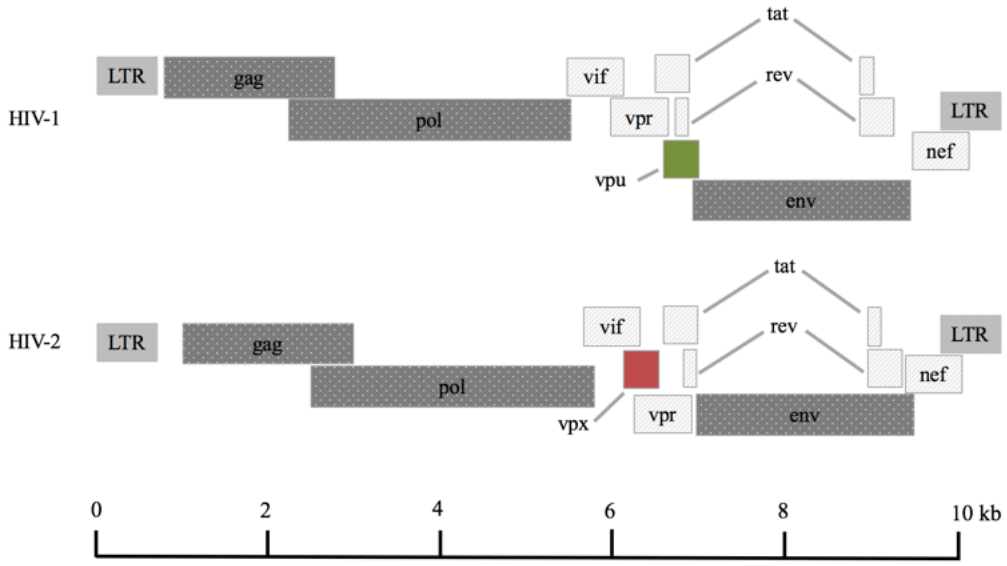

Figure 1.0.1. Genomic organization of HIV-1 and HIV-2. In color, are the two accessory proteins that are dissimilar to the human retrovirus: the HIV-1 vpu in green and the HIV-2 vpx in red.

is more closely related to simian immunodeficiency virus (SIV) with which it shares about 75% sequence homology (5). This is unsurprising given that HIV-1 subtypes are the result of zoonotic transfers of SIV from chimpanzees in the case of HIV-1 groups M and N and gorillas in the case of groups P and O, while HIV-2 is the result of zoonotic jumps of SIV from sooty mangabeys into human hosts (6,7).

Distribution

HIV-1 was the first of the two retroviruses to be isolated and is responsible for the global AIDS pandemic (8). Compared to HIV-1, the distribution of HIV-2 is far more limited: it is endemic in West Africa where sooty mangabeys roam (Figure 2). Evolutionary tree analyses of SIV_{smm} strains isolated from wild-living sooty mangabeys in Côte d'Ivoire and HIV-2 strains, revealed that the two main HIV-2 groups, A and B, are the result of two independent transmission events (Figure 3) (9). There are an additional six groups of HIV-2 (C to H), but these represent zoonotic transmissions with limited or no spread in the human population, with each transmission resulting in the infection of one or two humans.

As mentioned, HIV-2 is endemic in West Africa and in 2007 the World Health Organization reported that an estimated 1 to 2 million people were infected with the virus; an updated count is needed, especially given evidence showing a decrease in the prevalence of HIV-2 (10). Outside of West Africa, the majority of HIV-2 cases are in Portugal and France. The source of HIV-2 in these countries is different and related to their former colonial relationships: Portugal with Cabo Verde and Guinea-Bissau, and France with Côte d'Ivoire and Senegal (10). In the United States, the total number of HIV-2 infections is low. A total of 242 cases of HIV-2 were

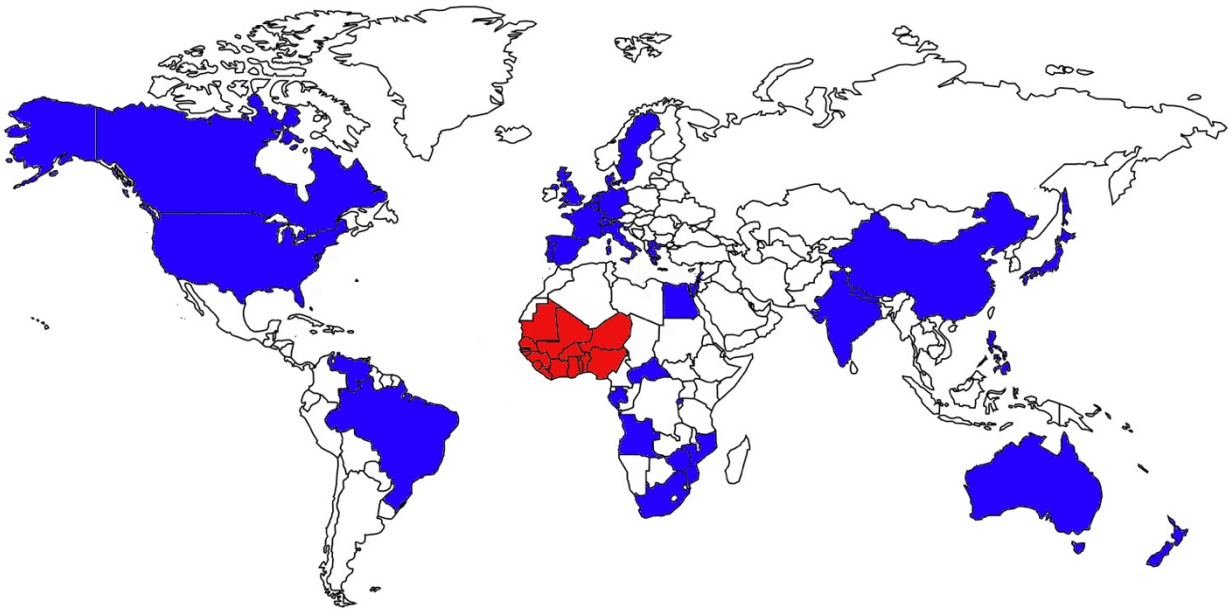

Figure 1.0.2. Global distribution of HIV-2. HIV-2 is endemic in west Africa (shown in red) and cases of HIV-2 have also been reported in a number of other countries (shown in blue), with the majority of HIV-2 cases outside west Africa occurring in Portugal and France.

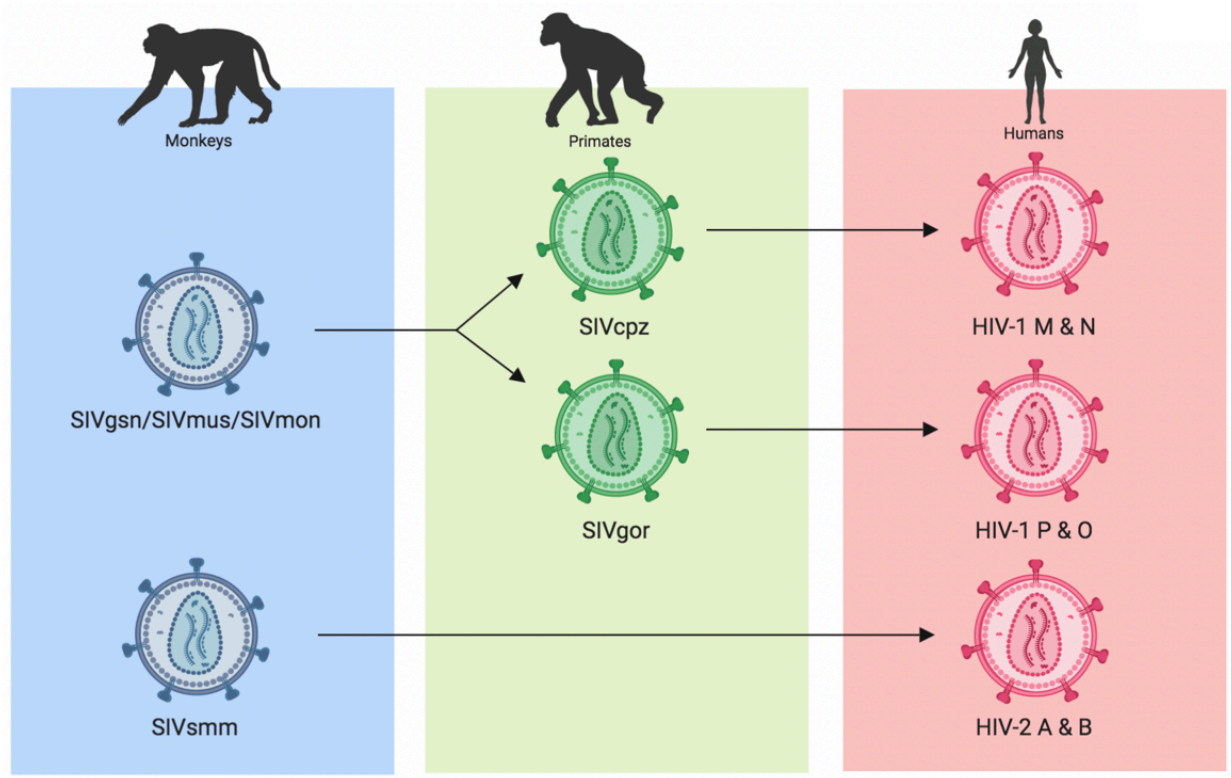

Figure 1.0.3. Origins of HIV-1 and HIV-2. Sooty mangabey monkeys (SMM) are a natural host of SIV, as are greater spot-nose monkey (GSN), mustached monkeys (MUS), and mona monkeys (MON). The journey SIV took to enter the human population and become HIV-1 had an intermediate ape step, with SIVcpz and SIVgor, while for HIV-2 the SIV virus jumped directly into humans.

reported to the Centers for Disease Control and Prevention (CDC) between 1988 and June 2010 (11). Only 166 of these fit the working definition of HIV-2 and cases were concentrated in the Northeast (66%), with the majority occurring in patients born in West Africa (81%) (11). As such, while HIV-1 in the United States is mostly maintained by new infections (38,739 in 2017), HIV-2 cases in the country can be described as imported (10,11).

HIV-2 Transmission

The dissimilar distributions of HIV-1 and HIV-2 are likely a factor of their different transmission efficiencies. The two viruses use the same routes of transmission: sexual contact, both vaginal and anal, directly into the bloodstream via needle sharing or a wound, and mother-to-child transmission (MTCT). Despite this, the transmission of HIV-2 is less efficient when compared to HIV-1 and these observed differences are related to the lower viral loads of HIV-2-infected individuals. Popper and colleagues compared the plasma viral loads of HIV-1 and HIV-2-infected antiretroviral therapy (ART) naïve study participants and found that the viral load is around 30 times lower in those infected with HIV-2 (12). Additionally, HIV-2 viral loads are undetectable in a large proportion of the HIV-2-infected population. Assessing viral loads in HIV-2-infected, ART-naïve individuals in Mali, Burkina Faso and the Côte d'Ivoire, Ekouévi and colleagues found that 46% of the study participants had viral loads below 10 copies per mL (13).

The comparatively lower viral loads of HIV-2 are not just found in the plasma, but also in other bodily fluids. In a study looking at levels of HIV RNA in infected Senegalese women, those infected with HIV-2 were less likely to have detectable levels of HIV RNA in cervicovaginal lavage (CVL) samples when compared to women infected with HIV-1: 58% versus 78% (14). Even when HIV RNA is detectable, lower levels of HIV RNA were also observed in HIV-2-

infected women (14). HIV-2, like HIV-1, is primarily spread by heterosexual transmission, with the level of HIV in the genital tract influencing the risk of transmission. The lower amounts of HIV-2 RNA found in the genital track are thus associated with the lower transmission efficiency of this virus.

Lastly, vertical or MTCT of HIV-2 is also lower compared to HIV-1, and in some countries, MTCT of HIV-2 is considered virtually nonexistent. In a prospective study of MTCT rates for HIV-1 and HIV-2 in Portugal between 1999 and 2005, the MTCT was 3.4% for HIV-1 and 1.5% for HIV-2 (15).

HIV-2 Pathogenesis

Transmission efficiency is not the only factor of HIV viruses that correlates with plasma viral load. Following a cohort of 180 HIV-1 male study participants, Mellors *et al*., showed that plasma viral load, rather than $CD4^+$ T cell counts, was the better predictor of progression to AIDS (16). Unsurprisingly, both viral load and pathogenicity are for HIV-2 compared to HIV-1 study participants (12). Both HIV-1 and HIV-2 infections are believed to follow a similar behavior in infected, ART-naïve individuals. The first stage is an acute infection, which in HIV-1 is marked by a decline in $CD4^+$ T cell counts and high levels of virus in the blood. Though there are no clinical data about acute infection in HIV-2 infected patients, it is believed that it takes on the same pattern observed for HIV-1 (17). This phase is followed by the asymptomatic phase where $CD4^+$ T cell counts recover somewhat and the virus reproduces at lower efficiencies. The third phase occurs after viral loads begin to increase, causing $CD4^+$ T cell counts to once again fall, progressing eventually to AIDS. Whereas HIV-2 infections can lead to all three phases, the majority of patients only experience phase 1 and 2; the percentage of HIV-2 patients who can be deemed LTNPs, that

is, individuals who maintain a $CD4^+$ T count greater than 500 cells/mm^3 without receiving ART, is as high as 95% (18). This is in stark contrast to HIV-1 infections, where less than 5% fit the clinical classification of LTNPs (19). Furthermore, for HIV-2 infected individuals who do progress to AIDS, the asymptomatic phase is longer: over ten years and up to decades compared to less than 10 years for those infected with HIV-1 (17,19).

The lower pathogenicity of HIV-2 was apparent soon after the retrovirus was discovered. In 1988, analysis of the *in vitro* cytopathicity of the HIV-2_{ST} strain isolated in Senegal, demonstrated no syncytium formation when co-cultured with uninfected $CD4^+$ cells and lower cell killing of phytohemagglutinin (PHA)-stimulated donor lymphocytes and SupT1 cells (20). The pathogenicity of HIV-2_{ST} is particularly diminished; the prototype group A strain HIV-2_{ROD}, on the other hand, induces both syncytia and cell killing, showing that even though HIV-2 strains are generally less pathogenic than HIV-1 strains, they still present varying ranges of pathogenicity (20).

Despite the differences in pathogenicity between the two retroviruses, and the observation that those infected with HIV-2 are more likely to be LTNPs, when HIV-2 infection does progress to AIDS the clinical features are similar to HIV-1-induced AIDS (17).

Thus, the generally held truths of HIV-2 infection are that it is less pathogenic than HIV-1 and the majority of those infected are LTNPs. A recent study, however, has called these assumptions into question. A long-term follow-up study in Guinea-Bissau found that a larger percentage of the HIV-2-infected study participants did in fact progress to AIDS when not placed on ART during the years they were followed, which varied from 4.9 to 16.2 years (21). Of the 87 HIV-2-infected participants enrolled 37, or 43%, progressed to AIDS. While 43% is still lower than HIV-1 (~95%), this is significantly higher than what has been previously reported for HIV-2

and bolsters the case for treating all HIV-2 infected individuals with ART (21). Interestingly, the Guinea-Bissau study included more male study participants than previous HIV-2 studies and earlier work from the Bissau HIV Cohort study group has suggested that HIV-infected males have a higher mortality than HIV-infected women independent of whether they were infected with HIV-1, HIV-2 or dually infected with both viruses (22). As such, the results from the follow-up study may reflect possible sex differences in mortality. If confirmed, these results complicate our previous understanding of HIV-2 infection and emphasize the need for better understanding of this retrovirus and its clinical outcomes.

HIV-2 Replication Cycle

While HIV-1 and HIV-2 may differ substantially in terms of distribution, transmission and pathogenicity, the replication cycle of the two viruses are exceptionally alike, though they diverge in a number of small ways. Here we'll go over the replication cycle of HIV-2, highlighting moments where the two viruses differ, and areas where more research or confirmatory experiments are needed (Figure 3).

Cell Tropism

HIV-2, like all viruses, needs to infect host cells to reproduce. Not all cells in the human body, however, are susceptible to HIV-2 infection. Similar to HIV-1, HIV-2 primarily targets cells of the immune system, specifically $CD4^+$ T cells, because these express the CD4 protein which mediates the initial interaction of the HIV virions with the soon-to-be infected cell. Primary macrophages can also be infected by HIV-2 isolates, though here the replication kinetics of HIV-2 differs from that of HIV-1, with a much higher burst of replication followed closely by latency

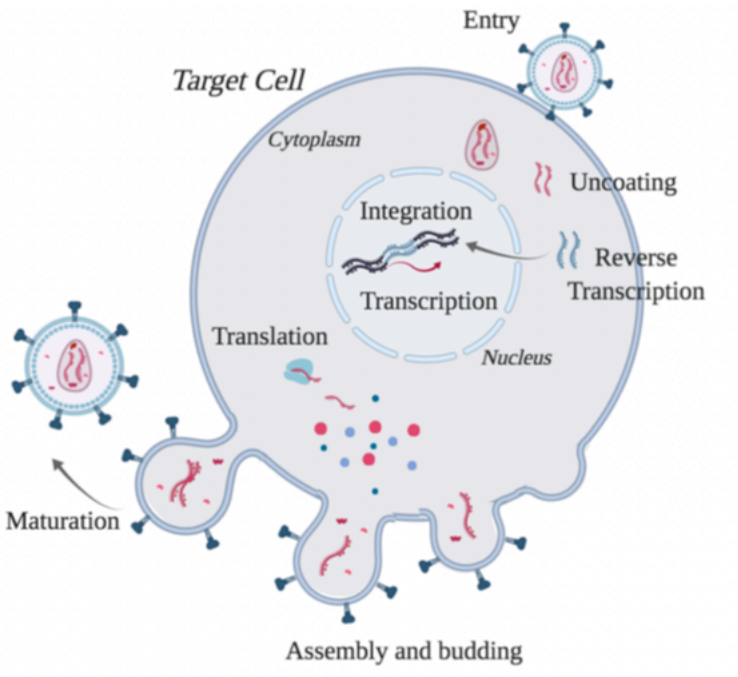

Figure 1.0.4. HIV-2 Replication Cycle. HIV-2 shares with HIV-1 all the classic steps of the retrovirus replication cycle, from entry to maturation post budding.

(23). Despite these other cell types, infection of $CD4^+$ T cells are the primary contributor to HIV-2 maintenance in infected humans.

Viral Entry

Viral entry begins with the binding of HIV-2 virions to specific cells in the host and is mediated by two glycoproteins on the HIV-2 envelope: gp125 and gp36, which assemble into trimers of non-covalent gp125-gp36 heterodimers with gp125 as the external glycoprotein and gp36 providing a transmembrane anchor (24,25). Entry requires an interaction between gp125 and the CD4 protein of the host cell (26). The HIV-2 Env is more closely related to the SIV Env, with which it shares a greater than 60% amino acid sequence homology, than to the HIV-1 Env (less than 40% homology) (27). The SIV, HIV-1 and HIV-2 Envs, however, are all structurally and functionally related, and all use CD4, with a few rare exceptions (28).

The interaction between the glycoproteins and CD4, however, is not enough to promote entry; similar to HIV-1, HIV-2 requires concomitant binding to a co-receptor to prompt the fusion of viral and cellular membranes (29,30). In HIV-1 entry, the most common co-receptors are the CXC chemokine receptor 4 (CXCR4) and the CC chemokine receptor 5 (CCR5) (31–33). These are also the two main co-receptors mediating HIV-2 virion entry, though CCR5 strains are more common in chronically infected asymptomatic HIV-2 patients (34). CXCR4 HIV-2 strains have only been isolated in patients with advanced disease and low $CD4^+$ T cell counts and, similar to HIV-1, are associated with a more rapid disease progression (34,35). Both HIV-1 and HIV-2 have been observed using alternative molecules for entry *in vitro*, with HIV-2 being the more promiscuous of the two retroviruses (36). CD4-independent entry has also been observed *in vitro* for both viruses, though more extensively so for HIV-2 where entry of CD4 negative cells can be mediated by CXCR4 or CCR5 (37,38).

There are five highly variable regions in gp125 (V1 to V5), and within the HIV-2 field there is a debate as to which regions are crucial for co-receptor binding (36). Work with chimeric viruses and binding to CCR5 or CXCR4 supports HIV-2 V3 as the crucial variable region, while site-directed mutagenesis of the V1/V2 region and CCR8, a non-canonical co-receptor, points to this region (24,36,39). As such, further research is necessary in order to tease out the regions of gp125 that are essential for HIV-2 co-receptor binding, not just to CCR5 and CXCR4, but also to the other alternative co-receptors such as CCR8.

Following binding of gp125 to CD4, conformational changes occur to the envelope glycoprotein causing it to reveal binding sites for co-receptors; the interaction between the virion and the co-receptors drives fusion. Examining dye transfer between dye-labeled HeLa cells expressing HIV-1 or HIV-2 envelope proteins and SubT1 cells, Gallo and colleagues observed that HIV-2 glycoprotein-mediated fusion is demonstrably quicker than HIV-1: 30 minutes instead of 60 minutes (40). The same research group observed that CD4-induced CXCR4 engagement was higher for HIV-2 gp125 than HIV-1 gp120, which they conclude is responsible for the more rapid fusion kinetics of HIV-2 envelope compared to HIV-1 envelope (40). Similar work has not been done with CCR5 strains. Despite this, these results, coupled with the observation of CD4-independent entry by certain HIV-2 strains, has been taken as a suggestion that the HIV-2 gp125 occasionally adopts a CD4-induced conformation even in the absence of binding to CD4 (34,38).

Reverse Transcription

Following entry, the conical core which contains two copies of genomic RNA and several virally encoded proteins, begins uncoating. Concurrently with this process, the genomic RNAs are reverse transcribed, forming the HIV double-stranded DNA molecules, which quickly become

associated with both viral and cellular proteins and are eventually transported into the nucleus. The uncoating process of HIV-2 has not been extensively explored. Recently, Takeda et al. used an *in situ* uncoating assay with fluorescent HIV-1 Vpr and HIV-2 Vpx – two accessory proteins found in the core of the mature virion – to compare the uncoating kinetics of the two retroviruses in the presence of Old World Monkey (OWM) or human TRIM5α, the tripartite motif protein isoform 5 alpha that is a significant restriction factor of HIV infection (41). They found that, in the presence of OWM TRIM5α, only one of the HIV-2 strains investigated showed accelerated uncoating like HIV-1, while in the absence of TRIM5α, the uncoating kinetics of HIV-2 are faster than those of HIV-1 (41). Further studies will be needed to uncover whether the same kinetics are observed in other HIV-1 and HIV-2 strains and to discern the biological basis for this difference.

Reverse transcription is catalyzed by the viral reverse transcriptase (RT). In HIV-2, the RT is a p68/p55 heterodimer that is more stable than the HIV-1 p66/p51 heterodimer, though the two share a similar domain structure in the subunits (42). The larger subunit has five domains, including the C-terminus RNaseH domain, which degrades the RNA strand in the RNA-DNA heteroduplex; the smaller subunit has four, tightly packed domains (42). Despite the structural similarities, there is one difference between the HIV-1 and HIV-2 RTs that impacts the efficiency of non-nucleoside RT inhibitors (NNRTIs). HIV-2 RTs are intrinsically resistant to NNRTIs; the crystallization and subsequent structure solution of the HIV-2 ROD RT by Ren and colleagues provides a potential mechanism for this resistance. The binding site of NNRTIs in the HIV-1 RT is in the palm domain of p66 and while an equivalent region is present in p68, there are alterations in the position of conserved residues as well as significant side-chain differences (42). These differences create a more constricted pocket that limits the binding of NNRTIs to the HIV-2 RT, thus making NNRTIs ineffective at restricting reverse transcription in HIV-2-infected cells (42).

For both HIV-1 and HIV-2 reverse transcription begins with the primer tRNA Lys3 interacting with the primer binding site (PBS) of the viral genome (43). Reverse transcription of HIV-2 can occur in cell types with dramatically different levels of cellular deoxynucleotide triphosphates (dNTPs), such as in active $CD4^+$ T cells but also in non-dividing cells like macrophages that have dNTP concentrations that are 50-200 times lower (44,45). In myeloid cells, the Sterile Alpha Motif (SAM) domain and Histidine-Aspartic (HD) domain-containing protein 1 (SAMHD1) depletes the dNTP pool, preventing efficient reverse transcription of HIV-1 (46). HIV-2, however, has a way around SAMHD1: the accessory protein X, also known as Vpx, which targets SAMHD1 for proteosomal degradation, allowing HIV-2 to replicate in macrophages and other non-dividing cells (47).

Given that HIV-2 strains have Vpx to counteract SAMHD1, Lenzi and colleagues explored the evolutionary effect of Vpx on the kinetics of the HIV-2 RT (47). Their results suggest that while HIV-1 and HIV-2 RTs have similar dNTP binding affinities, HIV-1 RTs have a higher rate of dNTP incorporation; furthermore, at low dNTP concentrations, the RTs of both retroviruses appear to have different DNA polymerase activity profiles (47). For Lenzi *et al.*, this suggests that HIV-1 has evolved a faster incorporation rate to make up for the lack of Vpx. Sakai and colleagues turned to phylogenetics to uncover insights into the relationship between RT and Vpx; they found that lentiviruses with anti-SAMHD1 activity clustered together in a lentivirus RT protein sequences tree and this cluster also contained viruses with low RT activity (48). They reasoned that the clustering suggested that RT and anti-SAMHD1 properties are linked, though cautioned that more sequences are needed to make a decisive conclusion on the matter.

Reverse transcription of the RNA genomes results in a DNA-RNA heteroduplex. The RNase H activity of RT degrades the RNA strand, exposing the single-stranded DNA molecule,

which then acts as a template for complementary DNA (cDNA) synthesis, resulting in the double-stranded viral DNA that is shuttled to the nucleus and integrated into the host genome. The RNase H activity of HIV-2 is ten times lower than that of HIV-1, though both enzymes have a similar cleavage pattern (49). This ten-fold difference is due to residue 294 in the small RT subunit, which regulates the affinity of RNase H for its substrate; WT HIV-2 has a glutamine in this position, but when this is mutated into a proline, a substantial increase in the RNase H activity of the HIV-2 RT is observed (49).

Nuclear Import

Following formation of the double stranded HIV DNA molecule, a preintegration complex (PIC) is formed involving various viral proteins such as integrase (IN) and the matrix protein (MA). In HIV-1, the PIC also includes Vpr, an accessory protein that has been shown to play a role in the nuclear import of the PIC (50). HIV-2 also carries Vpr and it displays a high degree of sequence homology to the HIV-1 Vpr [46,47]. Despite this, the HIV-2 Vpr has no role in nuclear import; a nuclear localization signal, however, was identified on amino acids 65 to 71 of the HIV-2 Vpx (53).

A number of host proteins are also associated with the PIC. Using a yeast two-hybrid screening method, Cheng *et al.* identified Hsp40, a cellular heat shock protein previously known to affect replication of numerous viruses, as a host factor that facilitates nuclear import in a Vpx-dependent manner (54). Hsp40 interacts with Vpx and its overexpression enhances nuclear import, while its down-regulation via siRNA decreased nuclear import (54). While the viral and host factors involved in HIV-2 PIC are being slowly elucidated, what actually happens at the nuclear membrane when the HIV-2 PIC arrives remains unknown.

Integration

Retroviral replication requires the integration of the double-stranded DNA into the host genome. This step involves the viral protein Integrase, which first binds to viral DNA at the long terminal repeat (LTR) and cleaves the 3'-end. Integrase subsequently promotes association with the target host genome and the strand transfer of the viral DNA into the host DNA. This leaves a small gap between the 5'-end of the viral DNA and the host DNA that is quickly repaired by the host cell's machinery.

Integrase contains three functional domains: an N-terminal domain, a catalytic core domain and a C-terminal domain. Mutational analysis of the HIV-2 integrase with a series of 41 site-directed point mutants found that most mutants equally affected integrase's ability to cleave the double-stranded viral DNA and strand transfer (55). Other mutations had more specific effects. Of the point mutations in the N-terminal region, for instance, two mostly affected DNA cleavage, one affected both DNA cleavage and strand transfer, and the others had minimal effect on integrase activity; these results suggest that the N-terminal, which contains a zinc-finger like motif, primarily functions in DNA cleavage (55).

The central region of integrase is the most conserved of the three among retroviruses and point mutations in two of the most conserved amino acids (Asp-116 and Glu-152), which are in the catalytic core, resulted in low levels or complete elimination of both DNA cleavage and strand transfer (55). Mutations in conserved amino acids in the middle domain saw various consequences, from no effect on integrase activity to effects on integration-site preference and for integration reactions at 1-hour but not 24-hours (55). Compared to the first two regions, the C-terminal region is less conserved and only one point mutation in this region led to a decrease in DNA cleavage and

strand transfer (55). These results are from *in vitro* studies and confirmatory *in vivo* explorations of the integrase domains are needed.

The HIV-2 integrase has a 53% amino acid sequence similarity to HIV-1 integrase; despite the difference, van Gent and colleagues found that the HIV-1 and HIV-2 integrases are equally capable of integrating both target HIV-1 and target HIV-2 DNAs (56).

In HIV-1, the integration site has an effect on the level of viral transcription and the same can be assumed for HIV-2. Work mapping 202 integration sites in primary peripheral blood mononuclear cells (PBMCs) following *in vitro* infection with the primary HIV-2 isolate p1629, found that HIV-2 has similar integration tendencies compared to HIV-1, as well as SIV (57). Specifically, HIV-2 has a preference for integrating into transcriptional units, a preference it shares with HIV-1 and SIV, though HIV-2 seems to singularly favor integration in the reverse direction (57). These results, of course, need to be confirmed *in vivo*. Another difference between the two retroviruses is that in infected participants, proviral integration into heterochromatin, which can block transcription, was more commonly found in HIV-2 participants (5/23) than in HIV-1 participants (0/21) (57).

Transcription

Transcription of the integrated HIV genome begins with the binding of both cellular and viral proteins to the 5'-LTR. While the cellular machinery is sufficient for limited viral transcription, viral proteins, particularly Tat, are necessary for productive expression. The 5' LTRs of HIV-1 and HIV-2 are both subdivided into U3, R and U5 regions, and Tat binds to the TAR element, which is located in the R region for both viruses. Despite these similarities, the TAR element of the two retroviruses is dissimilar: for HIV-1 it takes the form of a single hairpin, while

for HIV-2 there are two stem-loops (58). Both stem loops of HIV-2 are necessary to achieve high levels of transcription, though mutational analysis suggests that the first stem loop is more active than the second loop (59).

Upstream of the TAR element, is the enhancer region, which binds to a series of transcription factors. The binding sites of HIV-1 and HIV-2 are also substantially different (Figure 4). While the majority of HIV-1 viruses typically have two NF-kB, three SP-1, two AP-1 and one Ets binding sites (though there are variations between subtypes), HIV-2 viruses have no AP-1 or Ets binding sites (they contain a single peri-Ets, or pets, site instead), as well as a single working NF-kB, with the second NF-kB site mutated beyond functionality (60,61). In the LTR of both viruses, there are additional enhancers such as Pub1 and Pub2 in HIV-2 and NFAT in HIV-1 (62). It's possible that the divergent enhancers of HIV-1 and HIV-2 alter transcription efficiencies, thus explaining the differences in viral load observed in infected patients. Using a CAT reporter gene, Tong-Starksen and colleagues investigated how the HIV-1 and HIV-2 LTRs differed in their response to T cell activation signals and found that HIV-1 responds more promptly to T cell activation than HIV-2, suggesting that *in viv*o the enhancers possibly play a role in the different clinical outcomes of HIV-1 and HIV-2 infection (61).

Alternative splicing of the RNA transcripts allows HIV-1 to take full advantage of its genome, producing a series of transcripts that fulfill different functions such as RNA genomes for new virions and producing all the necessary proteins for continuing the replication cycle of the virus. While HIV-1 splicing has been exhaustively studied, the same cannot be said for HIV-2. Splicing sites do exist in HIV-2 and future research should be done to determine whether the ratios of different groups of spliced RNA transcripts are similar to those observed in HIV-1 infected cells.

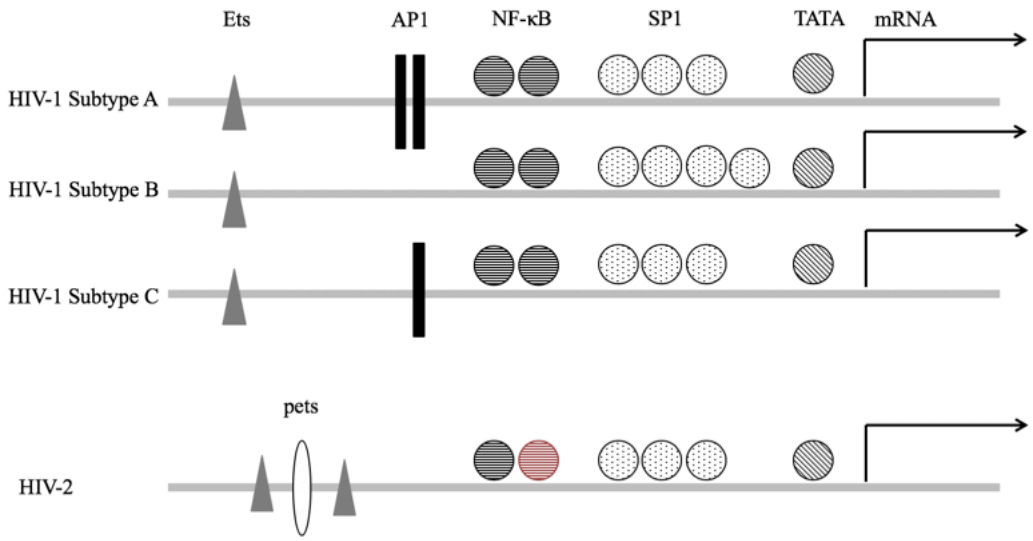

Figure 1.0.5. The binding sites in the enhancer regions of three HIV-1 subtypes and HIV-2. While the binding sites vary amongst HIV-1 subtypes, all contain one Ets site, two NF-kB sites, at least three SP1 sites and a TATA box. The HIV-2 enhancer region, on the other hand, has only one functional NF-kB site, with the second (in red) mutated beyond functionality, and instead of an Ets binding site it has a peri-Ets (pets) site.

Assembly and budding

The translation of HIV RNA transcripts uses host machinery to produce the proteins necessary for the assembly of new virions. While the efficiency of HIV-2 translation compared to HIV-1 has not been deeply explored, work with reporter constructs suggests that HIV-2 translation is less efficient and that this is due to the more complex structure of the HIV-2 TAR element (63). Furthermore, when the HIV-2 TAR element was deleted, there were increased levels of the reporter gene (63). This work only focused on genomic HIV-2 RNA, and it is possible that HIV-2 makes use of alternative splicing to overcome the TAR element and produce the proteins necessary for the assembly of new virions (64).

The final step in the retroviral replication cycle involves the assembly of new virions and their release from the cell. This step is essential for it is during assembly that all the viral components necessary for infectivity are gathered. One of the most crucial steps during assembly is the encapsidation of two copies of full-length genomic RNA, which are packaged as a linked dimer in both HIV-1 and HIV-2 virions. In the 5' untranslated region (UTR) of both viruses there is a dimerization signal that directs this process. In HIV-1 this signal is located in stem loop 1 (SL1), upstream of the major splice donor (MSD), and is known as the dimerization initiation site (DIS). In HIV-2, the exact location of the dimerization site is still debated. In the HIV-2 SL1, there is a palindrome that *in vitro* appears to be important for dimerization, though work *in vivo* has suggested that it is not (65). A second, more upstream palindrome has been proposed as the true dimerization signal (65).

During active HIV replication, many types of RNA transcripts are produced; for HIV-1, for instance, there are over 30 splicing alternatives. Unspliced genomic RNA represents only a small percentage of HIV RNA found in the cytoplasm (66). In contrast, genomic RNA constitutes

over 50% of the RNA in virions, which means that the encapsidation of the HIV RNA genome is selective (67). In HIV-1, this specificity is driven by both *cis*- and *trans*-acting elements: the *cis*-binding of packaging signals in the 5'-UTR and the *trans*-binding to the Gag protein. McCann and Levar searched for the HIV-2 equivalent of these packaging signals by investigating the effect of a series of deletions upstream and downstream of the HIV-2 splice donor (SD) (67). They found that while deletions downstream of SD had little effect on packaging, one of the upstream deletions (43 nucleotides from 392 to 434) had a profound effect on both packaging and replication (67). This deletion covers SL1, suggesting that SL1 plays a role in both dimerization and encapsidation.

Beyond packaging signals, HIV-2 has been proposed as having a very different packaging mechanism compared to HIV-1. This proposed mechanism is *cis*-acting, with Gag proteins packaging the RNA genomes that they are translated from. Work from Ni and colleagues, however, showed that packaging of genomic RNA shares many similarities between HIV-1 and HIV-2, adding evidence that HIV-2 may also be packaging in a *trans*-mediated manner (68).

Gag plays a major role in directing the assembly of new virions. In fact, with both HIV-1 and HIV-2, expression of Gag alone can lead to the production of virus-like particles (69). Assembly of the future HIV-1 virions occurs at the plasma membrane, and the shuttling of HIV-1 Gag to the membrane is dependent on clathrin adaptor proteins AP-1 and AP-3, which suggests that it requires vesicular transport. The N-terminal matrix (MA) domain of Gag is myristylated and, therefore, binds to membranes, adding further evidence for the vesicular transportation of Gag to the sites of assembly. The MA domain of HIV-2 Gag is also myristylated and the structure of the N-terminus is similar to that of HIV-1 (70). Using confocal immunofluorescence microscopy and siRNA-mediated depletion of adaptor protein complexes, Alford and colleagues found that HIV-2 Gag accumulates in intracellular compartments and particle production is dependent on the

adaptor proteins AP-3 and AP-5 (71). AP-5 localizes to late endosomal membrane, suggesting that endosomes may play a role in the assembly of HIV-2 but not HIV-1 particles.

During HIV-1 assembly, both Gag and Gag-Pol polyproteins are required to bind the RNA genomes and concentrate viral Env proteins at the assembly site, amongst other functions. As per usual, the assembly of HIV-2 is less understood, though believed to take the same steps as HIV-1 assembly. Interestingly, while the *env* sequence is generally known for its extreme heterogeneity, the gp41 subunits of HIV-1, HIV-2 and SIV *env* appear to contain a conserved domain involved in the formation of Env dimers and assembly (72). In fact, HIV-1 Env can form heterodimers with the Env proteins of both HIV-2 and SIV, though at a lower efficiency than homodimers (72). This coupled with other work, suggests that for all their differences, the assembly machinery of HIV-1 and HIV-2 is more similar than distinct.

Once all the components are assembled, budding occurs. At the same time, the viral protease is activated, cleaving the Gag polyprotein and leading to virion maturation. Protease is encoded in the *pol* gene and *in vivo* experiments investigating the maturation of HIV-2 virions showed that maturation only occurs when pol is expressed (73). With maturation, HIV-2 virions are ready to infect other susceptible cells and recommence the replication cycle.

Accessory Proteins

While certain cell types are more susceptible to HIV infection, this does not mean that they are completely devoid of protective mechanisms that help them counteract the invading retrovirus. Host restriction factors such as TRIM5α (affects efficient viral entry), tetherin (prevents budding of new virions) and members of the APOBEC3 family (promotes hypermutation of HIV) work to suppress viral replication (74).

At the same time, HIV viruses do not arrive at a cell unprepared; they carry in their small genomes the code for a number of accessory proteins that allow them to evade these restriction factors. Certain *in vitro* studies suggest that viral replication can occur without the presence of accessory proteins, but accessory proteins, both in HIV-1 and HIV-2, are necessary for efficient infection because they allow the viruses to escape host immune responses. This collection of accessory proteins differs slightly between HIV-1 and HIV-2. HIV-1 carries *vif* (APOBEC3), *vpr* and *vpu* (tetherin), as well as *nef*. HIV-2, on the other hand, encodes for *vif*, *vpr*, *nef* and *vpx* (17).

Vpu and Vpx do not have the same function: Vpu targets tetherin for degradation, while Vpx targets SAMHD1 for degradation, which restricts viral replication by depleting cellular dNTP levels, though SAMHD1 uses other means to limit HIV replication (46,47). The activity of SAMHD1 restricts HIV-1 infection in monocytes, dendritic cells (DCs), and macrophages, because reverse transcriptase is limited in cells with low levels of dNTP (46,75). HIV-1 can infect these cell types, however, if Vpx is supplied (76,77). While HIV-2 lacks a specific accessory protein for tetherin, it antagonizes it using the Env glycoprotein. This is a conserved function among HIV-2 primary isolates, though HIV-2 patients appear to harbor a mix of viruses that contain or lack the ability to antagonize tetherin (78).

In terms of structure, Vpx is similar to Vpr, a multitasking accessory protein whose most well studied activity is to arrest the cell-cycle in dividing cells at the G2 to M transition, though the reason for this in the viral life cycle is unknown (79). Both Vpr and Vpx have the same structure, a three-helix bundle, and the two genes are believed to be the result of a duplication or recombination (80). Of the two, the *vpr* gene is more widespread, found in all primate lineages, while *vpx* is found only in HIV-2/SIVsmm/SIVmac and SIVrcm/SIVmnd2 (80).

Looking only at the collection of accessory proteins and their functions alone, HIV-2 is the better equipped of the two viruses: able to counteract all the same host restriction factors as HIV-1, with the addition of having Vpx to tag SAMHD1 for degradation. The observed difference in pathogenicity between the two viruses is, therefore, unlikely to be due to accessory proteins.

The HIV-2 Reservoir

Recently, Saleh *et al.* raised the possibility that HIV-2 could provide novel strategies for a long sought out goal of the HIV field: a cure for HIV-1 (81). While ART can control HIV-1 infection, transforming what was once a death sentence into a manageable, chronic infection, treatment does not cure HIV-1 infection. The virus persists in a latent form in resting $CD4^+$ T cells, forming the long-lived and stable reservoir from which HIV can re-emerge if treatment is ceased or altered. As such, in order to cure an HIV-infected individual, the reservoir has to be purged or, alternatively, permanently locked.

The main intervention currently under investigation in the HIV-1 cure field, is latency reversal via pharmacologic agents, also known as "shock and kill." In this strategy, latency reversing agents (LRAs) "shock" or push the provirus out of latency, leading to the production of HIV-1 RNA and proteins, and the infected cell is killed either by the virus or by the immune system (82). If enough provirus-harboring cells are killed, the reservoir would be completely eliminated. LRAs "shock" infected cells through a variety of modes of action from histone deacetylase (HDAC) inhibitors to TLR agonists (83,84). Early work demonstrated that proviruses are often not induced during the first round of activation, requiring a second dose of the reactivating agent (85). Furthermore, while many reactivating agents are capable of increasing cell associated RNA and plasma viremia post-treatment, not all of them lead to a decrease in the reservoir size (86).To reach

high levels of reactivation, LRAs should be combined and work has been done to identify the most effective of these combinations (87).

Not every provirus in the reservoir has the ability to produce new virions when reactivated; only those that are replication competent hold this capacity. Work from multiple groups has shown that the HIV-1 proviral landscape is overwhelmingly defective, with large swaths of the provirus contingent containing large deletions and APOBEC-mediated G-to-A hypermutation, as well as other defects (85,88–90). Current work is focused on deciphering the nuances of the proviral landscape, particularly how it varies amongst patient populations and cellular subtypes.

Compared to HIV-1, the HIV-2 reservoir is less well comprehended. As previously shown, in treatment naïve study participants, infection with HIV-2 sees a lower plasma viral load compared to HIV-1 infection, despite both having comparable levels of proviral DNA (12). Saleh and colleagues reasoned that this suggests that HIV-2 infection has a post-integration block that affects viral replication and that this block could be exploited to functionally cure HIV-1 (81). Work in the 1990s showed that the HIV-1 and HIV-2 integrases are equally capable of integrating both HIV-1 and HIV-2 targets, and later work with *in vitro* infection of PBMCs revealed that the two human retroviruses have a preference for integrating into transcriptional units (56,57). The HIV-2 reservoir also appears to be a barrier to cure: in one case study out of Guinea-Bissau, HIV-2 plasma RNA levels rebound despite the HIV-1-HIV-2 dually infected patient being on therapy and previously having undetectable levels of both viruses for 19 months (91).

Until recently, the nature of the HIV-2 proviral landscape was largely unknown. Given that HIV-2 is more closely related to SIV than HIV-1, it was possible that the proviral landscape of this virus would look more like that of SIV-infected rhesus macaques, which is comparatively less defective, with at least 75% of the proviruses harboring some type of defect (92). In July 2019,

Bender and colleagues showed that this was not the case (93). The integrated genome of three HIV-2-infected individuals from Senegal was amplified and characterized, revealing that the HIV-2 proviral landscape is largely defective: only 1 of the 41 sequences generated was intact (93). Of the remaining 40 defective sequences, 66% contained large deletions and 12% had G-to-A hypermutation. The same paper also expanded on previous work from the same lab exploring the proviral landscape of suppressed, SIV-infected rhesus macaques; they found that in the seven macaques studied, 28% of the 287 sequences generated were intact, which is similar to what they had previously reported (92,93).

The analysis of the proviral landscape of the three HIV-2-infected individuals should be expanded and corroborated, but it does suggest that when it comes to molding the HIV integrated provirus, host factors, rather than viral factors, play a primary role. In this thesis, we build and expand on the work of Bender and colleagues, showing that their provirus landscapes results are consistent with our larger study population. Furthermore, we also look at how the HIV-2 reservoir behaves when pushed out of latency. For HIV-2 to provide new directions in the search for an HIV-1 cure, as Saleh *et al.* propose, we must first understand its reservoir. As such, a larger intent of this thesis, is to increase our current understanding of the HIV-2 reservoir and, in doing so, provide new directions for future HIV-2 work.

CHAPTER 2

HIV-2 Proviral Landscape Characterization

Abstract

The latent HIV reservoir is the major roadblock to cure because it is long-lasting, translationally silent, and therefore hidden from the immune system, reactivating when therapy is discontinued. Data at the case report level suggest that the HIV-2 reservoir is also a barrier to cure, with recrudescence of virion production when antiretroviral therapy is perturbed or changed. Whereas the HIV-1 reservoir and its reversal from latency have attracted considerable research interest, the nature of the latent HIV-2 reservoir is less well studied. To better understand the nature of the HIV-2 reservoir, we performed single genome amplification (SGA) and next-generation sequencing of near-full-length proviral sequences from nine participants with HIV-2 and characterized the integrated proviruses to determine what sequence defects they harbored. We found that the cumulative HIV-2 proviral landscape is 91 percent defective, similar to what we found in our comparator HIV-1 participant group (98% percent defective) and to what has previously been reported in the HIV-1 literature. Multiple types of HIV-2 provirus defects were identified and varied across participants. Among the types of defects found, HIV-1 and HIV-2 proviral sequences had similar proportions of proviruses harboring large deletions (61% and 65%, respectively) and G-to-A hypermutations (16% and 11%, respectively), though they differed in terms of the percent of proviruses with mutations and/or deletions in the PSI and MSD. Our results demonstrate that the HIV-2 proviral landscape is similar to that of HIV-1, suggesting that similar mechanisms may govern the generation of the lentiviral reservoir.

Introduction

In the quest for an HIV cure, the latent reservoir, defined as the cells that harbor transcriptionally silent yet replication competent integrated proviral genomes, is a major roadblock. The HIV-1 latent reservoir is established during acute infection and is heterogenous, infecting different T cell subsets and demonstrating different frequencies of intact and defective proviruses, among other variations (85,94,95). For the majority of those infected with HIV-1 who are on therapy, this latent reservoir represents a pool of cells harboring proviruses that can give rise to infectious virions should antiretroviral therapy be discontinued or become ineffective. With a mean T cell reservoir half-life of 44 months, the HIV reservoir will persist beyond the human life-span; an individual infected with the virus cannot be considered cured (96). As such, understanding the nature of the reservoir is considered essential to the development of a cure for HIV-1.

While there are still areas that require investigation, work from a number of groups has provided key insights into the nature of the HIV-1 reservoir. The majority of the HIV-1 reservoir is contained within memory $CD4^+$ T cells, even though the actual frequency of infected cells within this subset is low: an estimated one cell in one million (97). Sequencing-based approaches have shown that the majority of the HIV-1 reservoir is defective, with the majority of proviruses containing APOBEC-mediated G-to-A mutations and/or large deletions, as well as other defect types (85,88,90,98,99). The field is currently focused on the nuances of the HIV-1 reservoir, such as how it differs between patient cohorts and among the subpopulations of memory CD4 T cells.

Much less is known about the HIV-2 latent reservoir. One case report from Guinea-Bissau described a dually infected patient whose regimen change – from stavudine, lamivudine and lopinavir boosted with ritonavir (LPV/r) to once-daily didanosine, tenofovir and efavirenz – led to

detectable levels of plasma HIV-2 RNA and a decrease in $CD4^+$ cells (91). This case report demonstrated that, like HIV-1, HIV-2 can rebound when there are changes in the therapy regimen, thus suggesting that in those infected with HIV-2, the latent reservoir is also a barrier to cure (91). A greater understanding of the nature of the HIV-2 reservoir may also provide information about HIV latency more generally, generating novel insights to move us closer to a potential HIV cure (81).

HIV-2 is the less virulent, more globally restricted HIV that is found primarily in West Africa. While those infected with this lentivirus can also develop Acquired Immunodeficiency Syndrome (AIDS), progression is slower and the majority of those infected are considered long-term non-progressors, though a recent follow-up study in Guinea-Bissau from Esbjörnsson *et al.* has questioned these long-standing beliefs (18,19,21). Despite their divergent clinical outcomes, HIV-1 and HIV-2 share a similar replication cycle and a similar genetic organization. The two viruses only share about 40% overall nucleotide sequence homology, a factor of their different zoonotic origins: HIV-1 is a result of a transfer of SIV from chimpanzees and gorillas, while HIV-2 jumped into humans from sooty mangabey monkeys (5–7,9).

HIV-2 also shares the same routes of transmission as HIV-1, though transmission of this virus is less efficient, and the timing of the establishment of the HIV-2 reservoir has not been explored. The HIV-1 reservoir is established early in infection, though the exact timing is unknow. HIV-1 infection treated as early as ten days after the onset of symptoms still resulted in $CD4^+$ T cells with replication competent proviruses, and documented cases of infants treated within 30 hours of delivery showed that they experienced viral rebound following treatment interruption (94,100). Work in adult rhesus monkeys inoculated intrarectally with 500 $TCID_{50}$ SIVmac251explored the effect of early treatment on the establishment of the SIV reservoir and

found that it is also seeded early in infection and remains even when ART is commenced shortly after inoculation (101). In the case of the rhesus monkeys treated 3 days after inoculation, low levels of proviral DNA were detected in both lymph node mononuclear cells (LMNCs) and gastrointestinal mucosa mononuclear cells (GMMC), though not in PBMCs. Despite the lack of detection in PBMCs, viral rebound occurs when ART is discontinued showing that this persistent viral reservoir is established early in infection (101). The results from SIV-inoculated monkeys mirror those seen in HIV-1 participants where treatment during acute infection does not abrogate viral rebound during treatment interruption (102–104). As such, the prevailing hypothesis in the HIV-2 field is that the seeding dynamics of this retrovirus may be comparable to both SIV and HIV-1.

An intriguing similarity between HIV-1 and HIV-2 is that those infected with these viruses have similar levels of proviral DNA levels, suggesting that there is an as-of-yet unidentified immunological and/or viral mechanism that determines their different clinical outcomes post-integration (12). Furthermore, the longer clinically asymptomatic phase of HIV-2 and the higher number of long-term non-progressors implies that HIV-2 has a higher tendency towards less virus production and perhaps less virus transcriptional activity than HIV-1. This suggests that the mechanisms involved in the maintenance of latency may differ between the two viruses. An understanding of precisely how and why HIV-2 may have a greater propensity for transcriptional "silence" could lead to novel insights into how to approach the HIV-1 latent reservoir. To dissect these mechanisms, however, we need to first better understand the composition of the HIV-2 reservoir. Until recently, it was unknow whether the proviral landscape of HIV-2 would be more similar to that of HIV-1 or SIV. Bender and colleagues amplified 41 proviral sequences from three HIV-2 study participants on ART and observed that only one of these (2%) was intact (93). They

also looked at the reservoir of seven suppressed, SIV rhesus macaques and observed that 28% of the 287 proviral sequences are intact (93). Their work demonstrated that the HIV-2 proviral landscape is more similar to that of HIV-1 than SIV, but cautioned that a more complete picture of the HIV-2 reservoir was needed. In this chapter, we confirm Bender *et al.* finding and add more details to the HIV-2 reservoir, by characterizing the HIV-2 proviral landscape of nine HIV-2 study participants and comparing it to HIV-1 in samples obtained from 13 HIV-1 participants.

Material & Methods

Study population and sample collection

Cryopreserved peripheral blood mononuclear cells (PBMCs) were obtained from participants enrolled in the HIV Eradication and Latency (HEAL) cohort, a Boston-based longitudinal cohort of participants with HIV, or followed in the HIV clinics of Miriam Hospital in Providence (participants 101-104). A total of 13 participants with HIV-1 and 11 participants with HIV-2 were recruited. The clinical characteristics of the participants, including their current treatment regimens are in Table 1. Each participant had 90-180cc peripheral venipuncture samples collected in EDTA-containing tubes and samples were processed within 2 hours of collection. Plasma was separated following centrifugation at 1500 rpm for 10 minutes, and PBMCs were isolated by Ficoll-Histopaque density gradient centrifugation. Limitations on the volume of blood obtainable precluded the isolation of pure resting $CD4^+$ T cell populations. All participants gave informed consent prior to enrollment; the HEAL biorepository study was approved by the Partners Human Research Committee.

Amplification of Proviral DNA

Total DNA isolated from PBMCs for all the HEAL cohort participants was subjected to limiting dilution prior to near-full-genome length amplification with Platinum Taq HiFi

Table 1: Characteristics of Study Subjects

Patient ID no.	HIV	Sex	Age	Year of Diagnosis	Current CD4+ count, cells/mm^3	Current antiretroviral therapy*	Months plasma HIV RNA < 75 copies per mL
01	1	M	54	1998	606	3TC, ABC, EFV	24
04	1	M	37	2005	773	DRV/cobi, FTC TAF	>80
09	1	F	32	2010	1713	EVG/cobi, FTC, TDF	44
16	1	M	56	1984	267	ETR, FTC, RAL, TDF	10
19	1	M	54	2007	416	DRV/r, DTG, FTC, MVC, TAF	6
24	1	M	55	1996	691	EVG/cobi, FTC, TAF	12
25	1	M	31	2012	984	EVG/cobi, FTC, TAF	12
30	1	M	61	2006	1005	EFV, FTC, TDF	>27
31	2	F	55	2009	754	DTG, FTC, TAF	32
32	2	M	60	2005	330	DRV/cobi, EVG, FTC, TAF	>38
38	1	M	52	1996	695	DTG, FTC, TAF	>27
44	1	M	58	2001	448	DTG, FTC, TAF	16
50	1	M	65	~1990	899	DTG, FTC, RPV, TAF	42
52	2	M	67	2005	471	DTG, FTC, TAF	49
57	1	M	58	2012	1346	EVG/cobi, FTC, TAF	17
58	1	M	45	1994	647	DRV/cobi, FTC, TAF	29
61	2	F	44	2005	548	**	91
69	2	F	72	2005	607	DTG, FTC, TAF	>30
82	2	M	47	1995	1462	**	>66
84	2	F	61	1999	383	DRV/r, DTG, FTC, TDF	<12
101	2	F	44	2014	551	EVG/cobi, FTC, TAF	48
102	2	F	48	2002	1141	Treatment naïve	>110
103	2	M	59	2014	516	BIC, FTC, TAF	>12
104	2	M	60	2008	238	EVG,/cobi, FTC, TDF	>12

* 3TC, lamivudine; ABC, abacavir; BIC, bictegravir; DRV/cobi, cobicistat boosted darunavir; DRV/r, ritonavir boosted darunavir; DTG, dolutegravir; EFV, efavirenz; ETR, etravirine; EVG/cobi, cobicistat boosted elvitegravir; FTC, emtricitabine; MVC, maraviroc; RAL, raltegravir; RPV, rilpivirine; TAF, tenofovir alafenamide; TDF, tenofovir disoproxil fumarate
** Not on therapy

polymerase (Thermo Fisher Scientific). Both HIV-1 and HIV-2 were amplified using a previously published PCR protocol, but for HIV-2 the annealing temperatures were adjusted to 70 ºC for the first 10x rounds of amplification and 65 ºC for the following 25x rounds of amplification (105). The sequences of the HIV-1 primers for the first- and second-round PCR reactions were previously published and are as follows: first-round forward primer: 5'-AAATCTCTAGCAGTGGCGCCCGAACAG-3' (HXB2: 623-649); first-round reverse primer: 5'-TGAGGGATCTCTAGTTACCAGAGTC-3' (HXB2: 9662-9686); second-round forward primer: 5'-GCGCCCGAACAGGGACYTGAAARCGAAAG-3' (HXB2: 638-666); and second-round reverse primer: 5'-GCACTCAAGGCAAGCTTTATTGAGGCTTA-3' (HXB2: 9604-9632) [14,15]. The HIV-2 amplification primers were based on the HIV-2 prototype strain, ROD (GenBank: X05291.1), a subtype A virus that accounts for the majority of HIV-2 infections (106). The sequences for the first- and second-round PCR reactions were as follows: first-round forward primer: 5'- AGTGTGTGYTCCCATCTCTCCTAGTCGCCGCCTG-GTCAYT -3' (ROD: 240-284); first-round reverse primer: 5'-GAACACCCAGGCTCTACCTGCTASTGCTGGAGAGAACC -3' (ROD: 9537-9574); second-round forward primer: 5'- GRAAAATCCCTAGCAGRTTGGCGCCYGAACAGGGAC-3' (ROD: 285-320); and second-round reverse primer: 5'-GAGAACCTCCCAGGGCTCRATCTGCCAGCCTCTSCGCAGAGCGAC -3' (ROD: 9499-9543). PCR amplicons were directly sequenced by next-generation single-genome sequencing (NG-SGS) using an Illumina deep sequencing platform. For one HIV-2 participant (Heal 69), a different set of forward primers, designed based on HIV-2 subtype B prototype strain EHO (GenBank: U27200.1), were used after the original forward primers failed to amplify any sequences. These are: first-round primer: 5'-

TGCTTTGGGAAACCRARGCAGGAAAATCCCTAGC-3' (EHO: 817-850); and second-round primer: 5'- GCAGGAAAATCCCTAGCAGATTGGCGCCCGAA-CA-3' (EHO: 834-867); the same reverse primers were used.

Proviral Landscape Characterization Pipeline

Provirus amplicons were sequenced using Next-Generation Single Genome Sequencing (NG-SGS) using an Illumina deep sequencing platform by the Center for Computational and Integrative Biology (CCIB) DNA Core at Massachusetts General Hospital (MGH). Sequenced proviruses were manually analyzed and classified into one of five categories: large deletion (<8,000 bp); hypermutated; carrying a deletion in the packing signal (PSI in HIV-1, DM in HIV-2) or a mutation in the major splice donor (MSD in HIV-1, SD in HIV-2) site; carrying a large internal deletion (INDEL), carrying a stop codon or frameshift mutation; or intact using a pipeline adapted from previously published methods and are briefly summarized here (85,107). First, for each participant amplified proviruses with APOBEC3G-mediated hypermutation were identified using the Los Alamos National Laboratory HIV Sequence Database Hypermut algorithm in two steps (108). During the first step, proviruses were compared to HXB2 or ROD, noted as hypermutated or not, and non-hypermutated proviruses were used to create a consensus, non-hypermutated sequence. Then, all proviruses for each patient were compared to the patient-specific consensus sequence and determined as hypermutated using the Hypermut tool. Non-hypermutated sequences were then aligned to HXB2 or ROD using the Los Alamos National Laboratory HIV Sequence Database HIVAlign tool and INDELS were identified as sequences with an equal or greater than 2% difference in the length of each HIV gene (109). PSI deletions and MSD mutations were identified by aligning provirus sequences to HXB2 or ROD using Geneious Prime® 2019.1.1

software and checking for deletions and mutations in specific sites: the four stem loops of PSI and the MSD for HXB2, and the HIV-2 PSI and SD (65,110). Proviral sequences that lacked any of these defects were considered intact and labeled as such.

Results

Overview of Study Population

Eleven study participants with HIV-2 were recruited for this study, including eight who are on ART. An additional 13 participants with HIV-1 were also recruited. For HIV-1 participants, the median CD4 count was 625 cells/mm^3 (SEM=109), all patients were on ART, and all had plasma virus loads that were below the limit of detection. At the time of blood draw, eight of the eleven HIV-2 participants were receiving ART, with a median CD4 count of 548 cells/mm^3 (SEM=110) and nine of the eleven participants had undetectable viral loads. Given the relative rarity of HIV-2 participants, the HIV-2 and HIV-1 cohorts differ in regards to sex, CD4 counts, and months with undetectable plasma viral loads (11). In some cases, this may reflect differences in the biology of the two viruses and how they behave in human populations. Despite this, the HIV-2 and HIV-1 study populations have similar age distributions (Table 2).

We measured levels of cell associated RNA (caRNA), total DNA and plasma RNA for the 13 HIV-1 and 11 HIV-2 participants (Table 3). Levels of both total DNA and caRNA were higher in HIV-2 study participants, though for neither was this difference statistically significant. We also observed substantial variation between HIV-2 study participants for both measurements, as illustrated by the large interquartile range. For the 11 HIV-2 study participants, total DNA measures are 75 copies at Q1 and 1586 copies at Q3. For caRNA, Q1 is 105 and Q3

Table 2. Demographic characteristics of HIV-1 and HIV-2 Study Participants

	HIV-1 (N=13)	HIV-2 (N=11)	HIV-2 on ART (N=8)	HIV-2 not on ART (N=3)
Male, n (%)	92.3%	45.5%	50%	33%
Age, median, years	54	59	60	47
Years since diagnosis, median	18	14	14	17
Months undetectable viral load, median	24	38	31	91
CD4 count, median	695	548	494	1141
Race				
White	46.2%	0	0	0
Black	38.5%	81.8%	87.5%	66.6%
Hispanic	15.4%	0	0	0
Asian	0%	9.1%	0	33.3%
Biracial	0%	9.1%	12.5%	0

Table 3. Participant Reservoir Measurements

	Age (MEAN± SEM)	CD4 count cells/mm^3 (MEAN± SEM)	HIV DNA Copies per 10^6 PBMCs				HIV caRNA Copies per 10^6 PBMCs				Plasma HIV RNA Copies per mL
			Q1	Median	Mean	Q3	Q1	Median	Mean	Q3	
HIV-1	50 ± 4	806.9 ± 109.1	64	92	155	199	286	737	1626	1512	Below LOQ[1]
HIV-2	55 ± 3	636.5 ± 109.7	75	122	1496	1586	105	3126	3706	7414	Below LOQ[1,2]
HIV-2 on ART	59 ± 4	554.6 ± 89.4	82	241	1840	3340	45	2256	3271	7223	Below LOQ[1,2]
HIV-2 not on ART	46 ± 1	1050 ± 267.7	65	85	579	1586	3108	4076	4866	7414	Below LOQ[1,2]

[1]Measured supernatants were either undetectable or detectable but below the limit of quantification (LOQ) of 20 copies per mL

[2] For two HIV-2 participants, the measured plasma RNA copy was above the LOQ. HEAL 82 (not on ART) had 39 copies per mL, and Miriam 101 (on ART) had 48 copies per mL

7414. We did not observe the same level of variation in terms of total DNA and caRNA in the HIV-1-infected study participants. Higher levels of total DNA in HIV-2 study participants suggests a longer seeding period, potentially due to challenges in diagnosing HIV-2 infection, the lack of a quantitative HIV-2 RNA assay for assessing treatment effectiveness and the absence of clear guidelines for treating patients infected with HIV-2 (111).

Comparing the proviral landscape of HIV-1 and HIV-2

We amplified and sequenced near-full-genome length proviral amplicons for all 13 HIV-1 participants and nine of the eleven HIV-2 participants, including five of the eight participants on ART (Figure 2.01). Two participants – 102 and 103, recruited from Miriam Hospital – failed to produce amplicons with either the ROD or EHO primers and were removed from the participant cohort and all subsequent results do not include them. We set an endpoint of 20 near-full length sequences per participant for the SGA step, which we defined as sequences of greater than 8,000 bp in length. Twenty or more near-full length amplicons were achieved for all but one HIV-1 participant (Heal 09). We also counted the total number of amplicons generated with less than 8,000 bp and these were also sequenced. In total, we generated 257 and 235 near full-length amplicons for HIV-1 and HIV-2, respectively. There were an additional 436 and 443 amplicons with large deletions (<8,000 bp) for HIV-1 and HIV-2, respectively. The average total proviral genomes amplified for each participant was 53 and 68 for HIV-1 and HIV-2 study participants, respectively.

HIV-1 and HIV-2 amplified proviruses were deep sequenced using NG-SGS Illumina sequencing and analyzed to see if they were defective or intact. Defective sequences were further categorized as containing a large deletion, hypermutated, having a deletion in the packaging signal

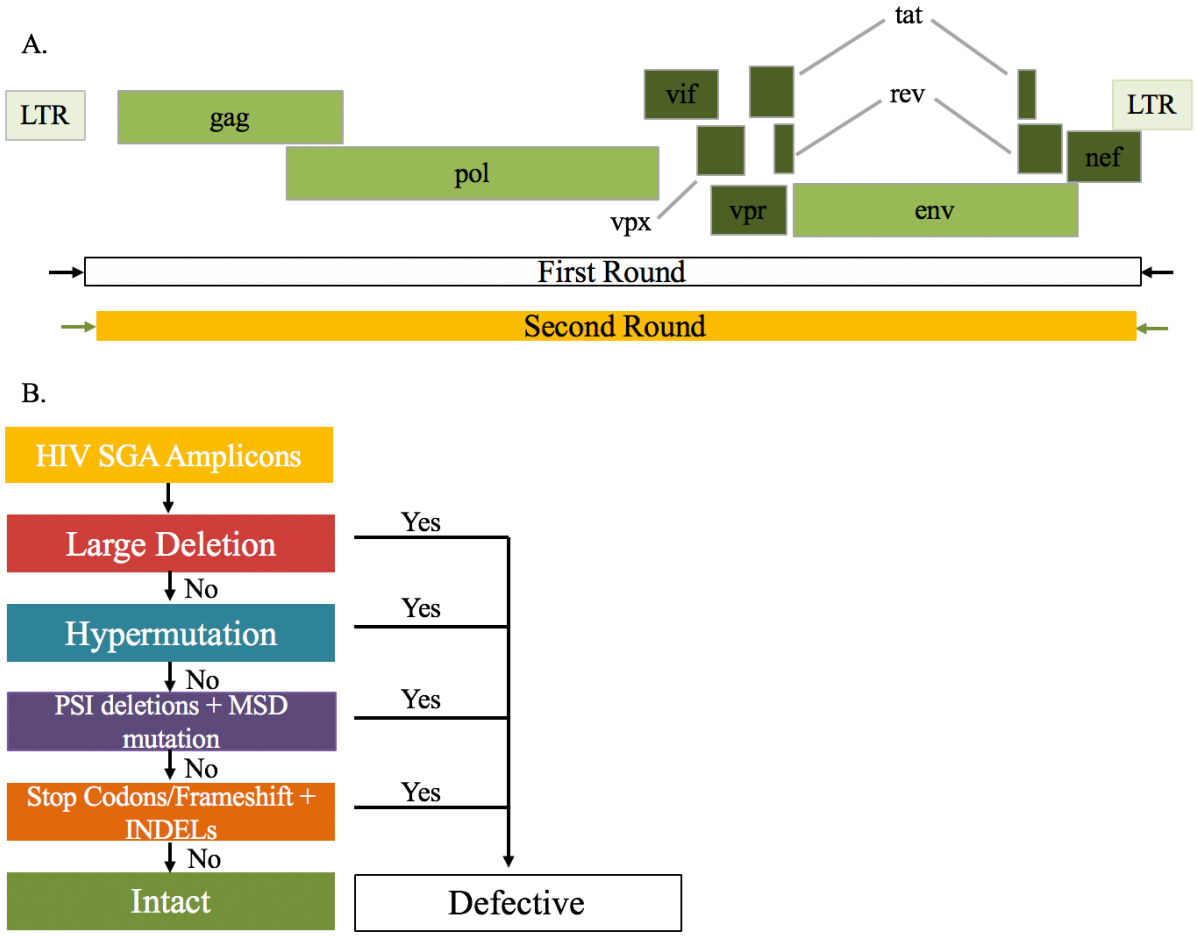

Figure 2.0.1. Amplification and characterization of near full-length amplicons of the HIV provirus. (A) The HIV-2 DNA genome with the first- and second-round single genome amplification (SGA) primers in the 5' and 3' LTR regions used to generate an amplicon that is directly sequenced using Next-Generation Sequencing (NGS). (B) Sequenced amplicons were passed through a pipeline to discern if they were intact or defective and, if defective, what type of defect they contain.

and/or a mutation in the major splice donor, carrying an internal deletion (INDEL) or a lethal mutation such as a frameshift or stop codon. The cumulative proviral landscape of our 13 HIV-1 participants and nine HIV-2 participants is shown in Figure 2.2. The individual proviral landscape of each of the nine HIV-2 and 13 HIV-1 participants is shown in Figure 2.3.

For both HIV-1 and HIV-2 study participants, defective proviral genomes constituted the majority of the amplified HIV reservoir: 98 percent for HIV-1 and 91 percent for HIV-2. Among the defective proviruses, HIV-1 and HIV-2 study participants had similar proportions of the various defect categories, with the exception of PSI deletions and MSD mutations (called DM deletions and SD mutations for HIV-2) where 15 percent of amplified HIV-1 proviruses contain such a defect, but only 8 percent of amplified HIV-2 proviruses have a deletion in this region.

While, Figure 2.2 provides information about the proviral landscape for our nine HIV-2 participants and allows us to compare it to our 13 HIV-1 participants, it masks the amount of variation observed between both HIV-1 and HIV-2 infected study participants in terms of their individual proviral landscapes. Figure 2.03 illustrates the participant-level variation in proviral reservoir composition, with some participants showing no intact proviruses among the sequences amplified (e.g., Participant 01 and 32) and others showing a much larger percentage of intact amplicons (e.g., Participant 04, 61 and, most dramatically, 31). This variation has an impact on the cumulative proviral landscapes generated for both HIV-1 and HIV-2 study participants. While we observed a greater proportion of intact proviral sequences in HIV-2 participants, relative to HIV-1, this difference was driven by a large proportion of intact sequences isolated from one participant with HIV-2. For participant 31, 30 out of 33 amplicons (91%), were intact by our analysis. This drove the cumulative intact percentage of HIV-2 proviral sequences to 8.7 percent. If the sequences

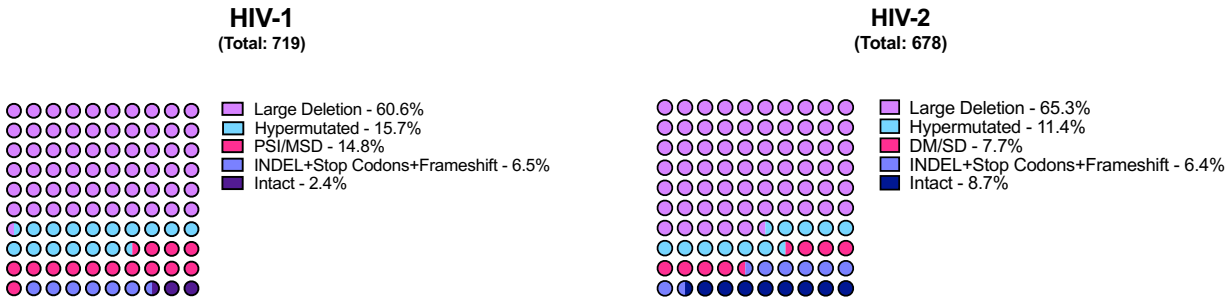

Figure 2.0.2. The cumulative proviral landscapes of HIV-1 and HIV-2 study cohorts. The cumulative percentage of each of the five categories is shown for the thirteen HIV-1 participants (left) and nine HIV-2 participants (left). Color key: Large deletion (lavender); hypermutated (sky blue); PSI deletion and/or MSD mutation (magenta); INDEL, stop codons and/or frameshift mutation (orchid); and intact (eggplant).

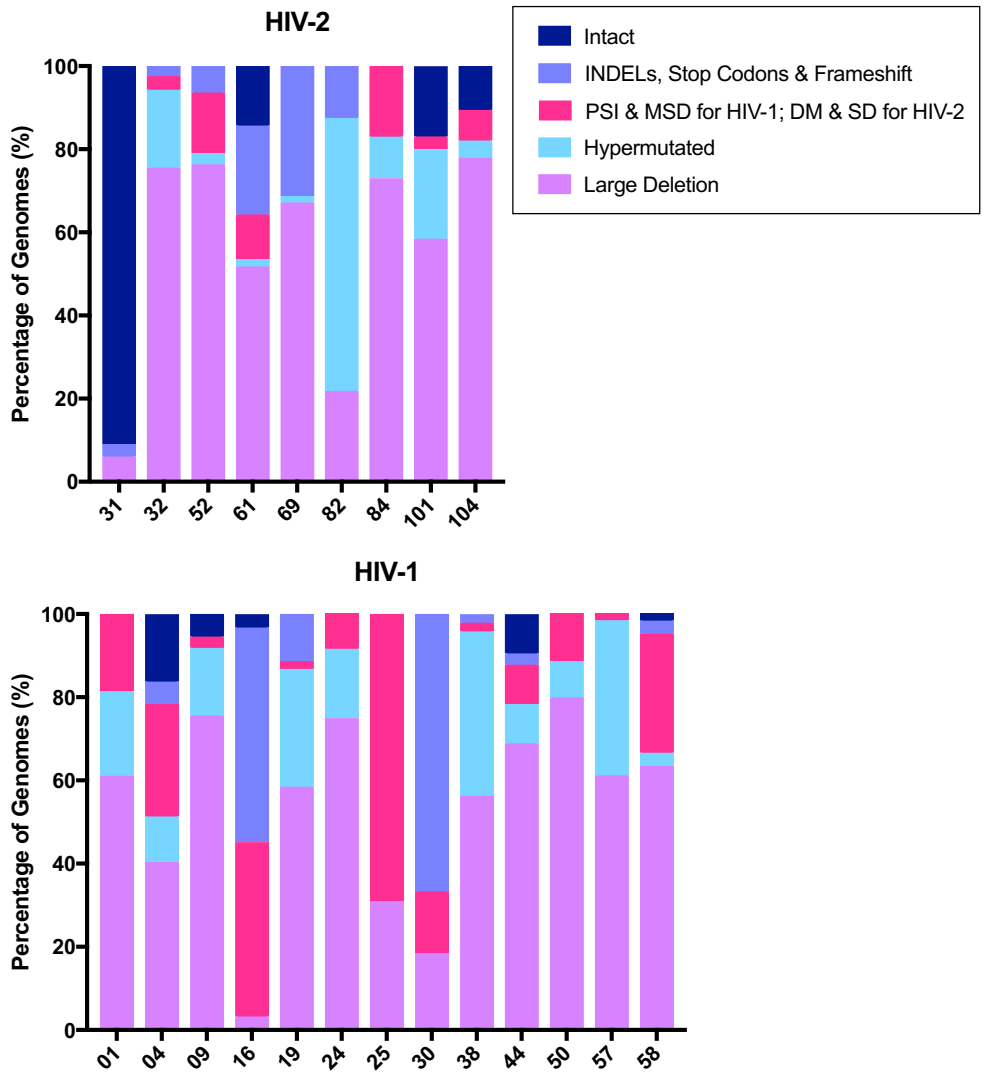

Figure 2.0.3. Comparison of individual proviral landscapes for HIV-2 and HIV-1 participant cohorts. Each study participant, shown on the x-axes, is represented by a column subdivided into the percentage of genomes in each category. Color key: Large deletion (lavender); hypermutated (sky blue); PSI deletion and/or MSD mutation (magenta); INDEL, stop codons and/or frameshift mutation (orchid); and intact (eggplant).

from participant 31 are excluded, the percent of intact sequences decreased to 4.5% for all HIV-2 participants.

Of the defect categories included in our pipeline, large deletions were the most common defect for both HIV-1 and HIV-2 participant cohorts, with 61 and 65% of genomes, respectively, carrying large deletions. For both HIV-1 and HIV-2, deletions covered a large percentage of the genome, an average of 64.8% for HIV-1 and 41.4% for HIV-2; this difference was significant ($p<0.0001$). The two viruses also differ in the distribution of the deletions, with HIV-1 having more deletions in the 3'-end (67%), while for HIV-2 there were a higher percentage of deletions in the 5'-end (62%).

Evidence of Clonal Expansion in HIV-2

To evaluate for cross-contamination between HIV-2 samples in the amplification process, we performed a phylogenetic analysis and generated a neighbor joining tree that included all near full-length sequences obtained. A total of 158 sequences were included; hypermutated sequences were not included in the phylogenetic analysis (Figure 2.5). For each participant, distinct branches were observed, ensuring that there was no cross-contamination between participants during any step of the SGA procedure.

While integration site analysis was beyond the scope of this initial work, for certain participants there is evidence of clonal expansion in the phylogenetic tree (Figure 2.5). Amplification of the provirus of participant 31 generated 25 amplicons of the same length (9,304 bp), of which, 10 amplicons and 11 amplicons form two clusters that differ due to a single mismatch at position 9,303. The other four amplicons of the same length have other mismatches.

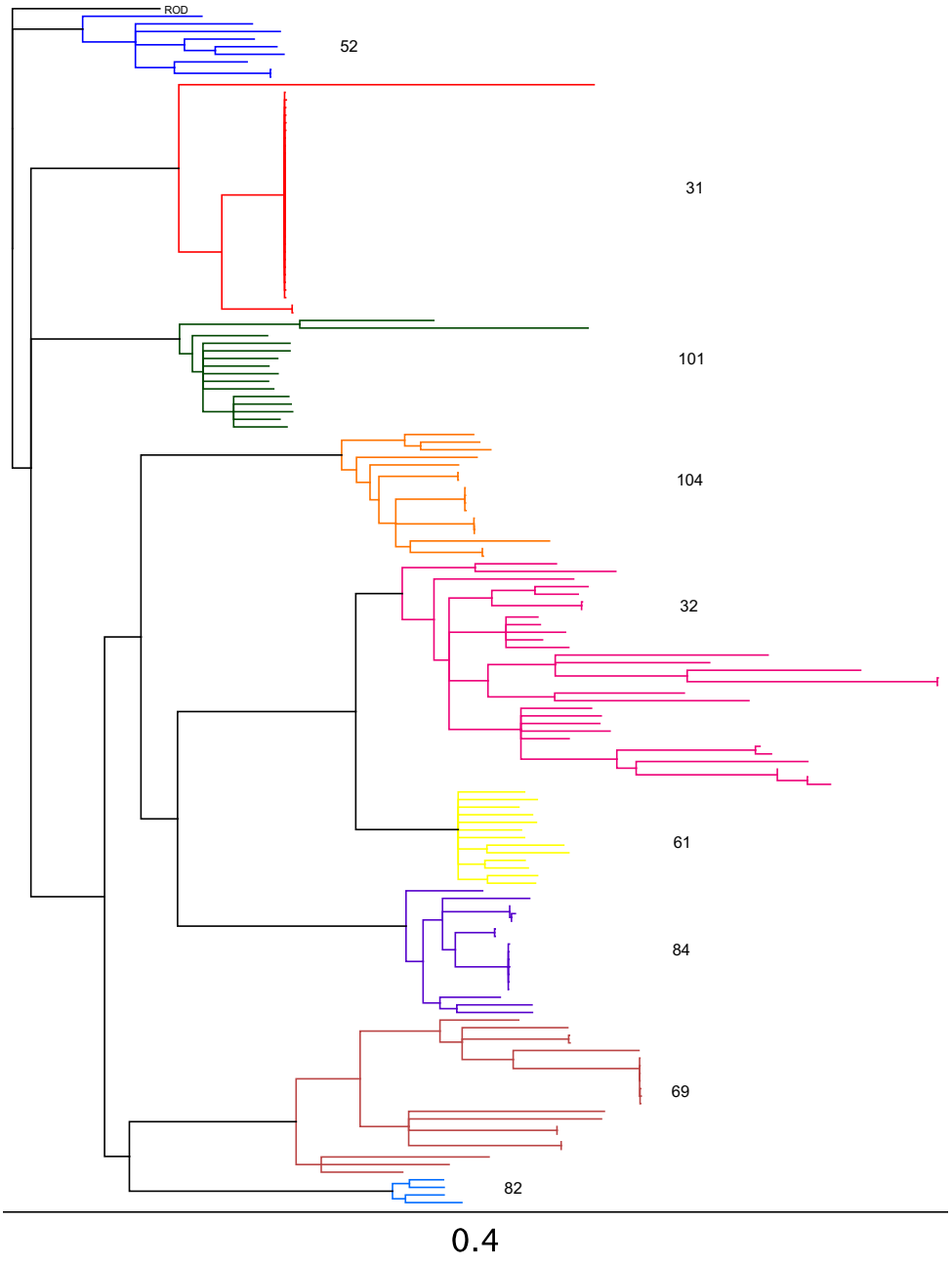

Figure 2.0.4. Neighbor-joining tree for all non-hypermutated HIV-2 near-full-length proviral sequences. The table includes all nine HIV-2-infected study participants as well as ROD, the HIV-2 reference strain, as an outgroup. Each participant is color-coded and the ID of each participant is on the right of the tree.

Two or more identical proviral amplicons are also observed in participants 69, 84 and 104.

Discussion

Until recently, the nature of the HIV-2 proviral landscape was largely undefined. Bender and colleagues amplified 41 sequences from three HIV-2 participants on ART and found that 40 of them (98%) were defective, showing that the HIV-2 proviral landscape more closely resembles that of HIV-1, which is largely defective, than SIV (85,88,89,92,93,99). Our amplification of 678 proviral genomes, 235 of which are near-full-length, from nine HIV-2 study participants, expands and adds to their work. We also amplified 693 proviral genomes, including 257 which are near-full-length, from 13 HIV-1 participants and analyzed them using a same pipeline. Our analysis found that the HIV-2 proviral landscape of our nine study participants was approximately 91 percent defective, a proportion more similar to that of HIV-1 than SIV, which Bender *et al*. found to be 28% intact (93). The types of defects we identified in the HIV-1 and HIV-2 proviral landscapes were also similar, particularly in terms of the percentage of total sequences categorized as containing a large deletion (61 percent and 65 percent, respectively) or APOBEC-mediated hypermutation (16 percent and 11 percent, respectively). The HIV-2 percentages of these two defect categories are also similar to what Bender *et al.* found for their three HIV-2 study participants: 66% of genomes had large deletions and 12% had hypermutations. For the SIV proviral landscape, on the other hand, 58% of genomes had large deletions, 13% were hypermutated and another 19% had both large deletions and hypermutation (93). Our investigation of the HIV-2 proviral landscape coupled with Bender *et al.*, suggests that the HIV-2 proviral landscape is more similar to HIV-1 in humans than SIV in non-human primates, despite HIV-2 being more closely related to SIV and sharing a greater sequence homology with this virus (5).

Beyond a general similarity in the proportions of defective proviruses, HIV-1 and HIV-2 reservoirs share other resemblances. Our results show that for the participants with HIV-2 we studied, the character of the provirus varies greatly amongst individual patients. For some, intact proviruses make up a larger proportion of the total amplicons (as seen in participants 31 and 61). On the other hand, and similar to what has been reported for other study cohorts as well as our participants with HIV-1, for a number of participants no intact proviruses are amplified, suggesting that for many individuals with HIV-2, the percentage of intact proviruses may be below the average of 9 percent we found. In fact, when participant 31 was removed from the analysis, the proportion of intact amplicons for the remaining eight participants with HIV-2 fell to 5%, which is much closer to what we found for our HIV-1 participants (2.4 percent) and also what has been previously described for HIV-1 (85,88–90,98,99).

In their analysis of the proviral genomes of three HIV-2 participants, Bender *et al.* also looked for evidence of clonal expansion, and found a pair of identical sequences for one of the participants. In our expanded cohort, we identified two or more identical proviral amplicons for four HIV-2 participants: 31, 69, 84, and 104. Of these, 31 has the greatest number of identical sequences, with two clusters of ten and 11 identical sequences. Clonal expansion, which occurs in those infected in HIV-1, plays a role in maintaining the HIV-1 reservoir (90,98,112). Integration site analysis of the HIV-2 proviruses will be needed to confirm clonality of the identical sequences, but the evidence we found here in four of the 9 participants suggests that clonal expansion also plays a role in maintaining the HIV-2 reservoir.

The ways in which the HIV-2 proviruses were defective also varied amongst individual participants, and this is shared with the HIV-1 study participants. While large deletions constituted the bulk of the defects for many participants, in others G-to-A hypermutation, mutations and

deletions in the packaging signal and major splice donor, stop codons and frameshift mutations were the primary defect identified. This suggests that there are many roads to defective proviruses in the HIV reservoir. Pollack *et al.* have proposed that when defective HIV-1 proviruses are expressed, different types of defects may be targeted by HIV-1-specific cytotoxic T lymphocytes (CTLs), thus contributing to the overall nature of the proviral landscape (113). The HIV-2 reservoirs we investigated show similarly varied distributions in terms of the types of categories and the percentages of proviruses in each category; investigating whether defective HIV-2 proviruses are expressed and, subsequently, targeted by CTLs, will demonstrate where the HIV-2 proviral landscape is also shaped by similar forces.

The one defect for which there was a greater disparity in the total percentage of sequences between the HIV-1- and HIV-2-infected participants were deletions in the PSI and mutations in the MSD: 14.7 and 7.7 percent, respectively. One possibility for this disparity is the difference in the size of the PSI, or packaging signal, of HIV-1 and HIV-2 viruses. The packaging signal is required for the correct packaging of the both HIV-2 and HIV-1 RNA genomes into assembling virions and in the HIV-1 5'-UTR, it is formed by four stem loops (SL1, SL2, SL3 and SL4) that together span a non-continuous area of almost 100 nucleotides (the MSD is in SL2) (107,114). The packaging signal of HIV-2, on the other hand, is much shorter: a 43 nucleotide stem loop called SL-1 upstream of the major splice donor (65,67). Given that the PSI of HIV-1 encompasses an area in the 5'-UTR that is two times longer than the HIV-2 PSI, this might explain the almost double increase in the percentage of proviruses that are deemed defective due to deletions in the PSI and mutations in the MSD.

A potential limitation of the study is our inability to amplify HIV-2 group B sequences. Our initial set of primers were based on ROD, the canonical sequence for HIV-2 group A, and this

set successfully amplified eight of the 11 HIV-2 participants. A second set of forward primers based on EHO, the canonical sequence for HIV-2 group B, was designed for the remaining three participants and successfully amplified participant 69. Two participants, 102 and 103, failed to amplify with either ROD or EHO primers, which may be due to a low total DNA or to primer mismatch. It's unknow whether the proviral landscape varies between HIV groups; the characterization of the proviral genomes of participant 69 suggests that this is not the case. Nonetheless, developing a set of primers that consistently amplifies HIV-2 group B proviruses is necessary.

Our investigation of the reservoir of nine participants with HIV-2 built on the previous work by Bender *et al.*, and together they reveal a proviral landscape that more closely matched that reported for HIV-1. Further research will delve into the nuances of this similarity, particularly in determining whether clonal expansion occurs with HIV-2. Though HIV-2 is more closely related to SIV, with which it shares a higher sequence homology, its provirus landscape mirrors that of HIV-1. This suggests that host factors, rather than viral factors, play a bigger role in the shaping of the HIV reservoir, both during the establishment of the reservoir and subsequently during its maintenance. Uncovering these factors and seeing how their effect on HIV-1 and HIV-2 diverges, may uncover new targets to be exploited in an HIV cure strategy.

CHAPTER 3

HIV-2 Reservoir Dynamics and Reactivation from Latency

Abstract

The HIV latent reservoir, a pool of replication competent integrated viruses that are not actively producing new HIV virions yet retain the ability to do so, is a major barrier to cure. Within the HIV-1 field, "shock and kill," where reactivation of latently-infected cells is followed by cell death, has been proposed as a potential cure strategy and has been investigated over the past decade. HIV-2, which infects 1 to 2 million individuals and is endemic in West Africa, has been less well studied. Furthermore, the reservoir dynamics of HIV-2 in virologically suppressed patients are unknown. To understand how latency reversing agents (LRAs) may induce HIV-2 transcription and virion production, we performed *ex vivo* reactivation experiments with peripheral blood mononuclear cells (PBMC) isolated from participants with HIV-2. When we control for total HIV DNA, we observed that HIV-1 and HIV-2 PBMC produce comparable amounts of cell-associated RNA following reactivation. However, at the supernatant RNA level, HIV-1 PBMCs produced significantly higher levels of virion-associated RNA, when compared to HIV-2 PBMCs following reactivation with bryostatin, romidepsin, PMA/ionomycin and αCD3/αCD28 beads. This finding suggests that LRA can reactivate HIV-2 transcription to produce cell-associated RNA but that a block exists which prevents virion production at rate equivalent to HIV-1.

Introduction

Acquired immunodeficiency syndrome (AIDS) can be caused by two lentiviruses: HIV-1, which is responsible for the global AIDS pandemic, and HIV-2, which is endemic in west Africa and is estimated to have infected 1 to 2 million individuals (115). While the two viruses have similar genetic structures, these belie marked differences in their clinical outcomes. Of the two, HIV-1 has a higher pathogenicity with the majority of people living with this virus progressing to

AIDS; HIV-2, on the other hand, has a much lower pathogenicity, with up to 95% of infected individuals reported as long-term non-progressors (LTNPs) (18,116). Compared to HIV-1, infection with HIV-2 is also generally associated with a longer asymptomatic stage and lower mortality (116).

Disease progression to AIDS is strongly correlated with plasma HIV RNA levels, or viral load. In antiretroviral therapy (ART) naïve individuals, those infected with HIV-1 have a viral load that is 1.5 to 2 $\log_{10}$ higher than those infected with HIV-2 (12). Furthermore, in clinical settings in west Africa, 25 to 40% of HIV-2 infected, ART naïve individuals have viral loads that are below the detection limits of current assays (115). Levels of cell-associated RNA are also higher in HIV-1 ART-naïve study participants compared to individuals with HIV-2 (117).

While both plasma RNA and cell-associated RNA (caRNA) levels were higher in HIV-1 than HIV-2 infected ART-naïve patients, total and integrated DNA levels were similar for individuals living with either virus (117). This suggests that the slower disease progression of HIV-2 compared to HIV-1 is not due to lower rates of viral integration into the human genome, but rather to unidentified post-integration factors. Moreover, the ability of HIV-2 virus to establish a pool of integrated viruses that show subsequent limited transcription, has led some to suggest that this retrovirus has a higher predisposition for establishing a latent infection compared to HIV-1 (117).

Previously published comparisons of HIV-1 and HIV-2 DNA and caRNA were performed in ART-naïve participants. Similar comparisons have not been performed for participants with HIV1 or HIV-2 infection on suppressive ART. With the growing number of patients with HIV-2 on ART and the suggestion that HIV-2 may have a higher tendency for latency, we investigated the reservoir dynamics of HIV-2 in patients receiving therapy who have achieved virologic

suppression. Moreover, while latency reversal has been extensively studied in HIV-1 patients with the use of latency reversal agents (LRAs), the same has not been true in HIV-2-infected individuals. Should "shock and kill," a two-step process that first uses LRAs to reactivate HIV in latently infected cells, allowing these cells to then be killed by the immune system or drugs, become a viable cure strategy for HIV-1, it is critical to investigate whether it could also be used as an intervention for those living with HIV-2 (82,118–120).

"Shock and kill" is one of a handful strategies for targeting viral reservoirs and reducing them to a level that results in either viral eradication or extended remission in the absence of ART. Current eradication strategies aim to reactivate latent HIV transcription in the presence of ART and then clear the reservoir, defined as cells that contain proviruses capable of expressing infectious virus, either through direct viral cytopathic effects or immune-mediated cytolytic T cell activity (121). Over 160 compounds with *in vitro* LRA activity, acting through a variety of mechanisms, have been identified, including histone deacetylase inhibitors (HDACi) like vorinostat and romidepsin, protein kinase C (PKC) agonists like a phorbol 12-myristate 13-acetate (PMA) and bryostatin, and Toll-like receptor (TLR) agonists like GS-9620 and CL413 (122). Despite the large number of potential LRAs identified and the clinical studies that have occurred since 2012, a LRA that is able to both increase HIV RNA transcription and lead to a decrease in the size of the reservoir has yet to be identified (122,123). For instance, histone deacetylase inhibitors (HDACi) have been studied in small human trials as HIV latency reversing agents (LRA) but their effects have been mixed. Single dose vorinostat stimulates *in vivo* intracellular HIV-1 RNA synthesis, but at clinically relevant concentrations reactivates <0.1% of proviruses to produce virions (83,124). Furthermore, multiple vorinostat doses did not result in detectable plasma viremia and lead to a loss of HIV-1 transcriptional reactivation (125,126). The more potent

HDACi romidepsin and panobinostat induced virion production and plasma viremia in some studies, but not in others (87,127–130).

Unlike HDACi, which increase HIV-1 transcription but do not induce immune activation, T cell receptor agonists (e.g. PHA/IL-2) or compounds that activate PKC and Ca^{2+} influx (e.g. phorbol 12-myristate 13-acetate [PMA] with ionomycin) generally cause immune activation and maximal HIV-1 reactivation (131). The degree to which LRA that do not activate T cells can induce HIV-1 outgrowth remains an open question (127,132–134). Virion production may or may not be necessary for cure but, at a minimum, the synthesis of virus proteins is an absolute requirement for any strategy that leverages humoral or cell-mediated immunity to eradicate HIV.

TLR agonists work as LRAs by activating the TLR innate immune system, which, through a variety of pathways depending on the exact TLR targeted and cell type, subsequently triggers proviral reactivation (122). Agonists against several TLRs, including TLR1/2, TLR7, TLR8 and TLR9, have been shown to induce HIV expression in latent cells (84,135–137). TLR7 agonists have also been tested in SIV-infected rhesus macaques; macaques orally treated with a GS-986, an analog of GS-9620, saw a transient immune activation following dosage as well as detectable plasma viremia (138).

Beyond TLR7, a handful of SIV research teams have investigated the effect of other LRAs on SIV-infected macaques on antiretroviral therapy, many with the intent of using nonhuman primate (NHP) models to see what occurs at the tissue level following reactivation (139–141). HIV-2 is more closely related to SIV, with which it shares 75% sequence homology, than to HIV-1 and, as such, latency reversal behavior in SIV provides a model of what may be observed for HIV-2 (5). Del Prete and colleagues infused romidepsin into virologically suppressed, SIVmac239-infected Indian rhesus macaques and observed a significant increase in plasma

viremia compared to macaques infused with saline. This effect, however, was transient and limited to when the macaques were receiving romidepsin infusions (140). The following year, Policicchio *et al.* observed no increases in plasma viremia following a single four-hour romidepsin perfusion into virologically suppressed, SIVsmmFTq-infected Indian rhesus macaques; multiple doses of romidepsin were needed to reactivate SIV *in vivo* (139). Work in pigtailed macaques, dually infected with SIVDeltaB670 and SIV17E-Fr and then virologically suppressed with a central nervous system-penetrant ART regimen consisting of tenofovir, darunavir, ritonavir and the integrase inhibitor L000870812, investigated the effect of ingenol-B (10 nmol/l), a PKC activator, on latently infected $CD4^+$ T cells isolated from the macaques and observed a significant increase in viral transcription (141). They then treated the macaques with increasing levels of ingenol-B (0.4 mg/kg/day for 30 days, then 0.6 mg/kg/day for 10 days) and found no changes in viral load. When ingenol-B (0.5 mg/kg/day) was combined with vorinostat (6 mg/kg subq. infusion 4 times), an HDAC inhibitor known to synergize with PKC agonists, for 10 days, plasma viral loads became detectable in one of the three pigtailed macaques (141). Reactivation work in NHPs reproduces results seen in the HIV-1 field, where (a) *ex vivo* transcriptional reactivation in isolated cells does not necessarily lead to *in vivo* increases in viral load, and (b) combinations of LRAs and/or repeated dosing may be necessary (87).

A number of factors, such as transcriptional regulation and chromatin remodeling, are involved in establishing and maintaining latency and the mode of action of LRAs reflect this multiplicity: HDAC inhibitors increase the acetylation of histone proteins making integrated proviruses more accessible to transcription factors, while PKC agonists like bryostatin work through the NF-κB and AP-1 signaling pathways (122,142,143). The absence of cellular transcription factors like NF-κB in the nuclei of resting CD4+ T cells is one such contributor to

the maintenance of latency and it has the potential to affect HIV-1 and HIV-2-infected cells differently (142). The HIV-2 enhancer region of the long terminal repeat (LTR) and the transcription factor binding sites it contains is substantively different from HIV-1 (61). Notably, the HIV-2 enhancer region lacks AP-1 binding sites and contains only a single functioning NF-κB binding site; plasmids containing the HIV-2 enhancer region were less responsive to T cell activation signals than plasmids containing the HIV-1 region (61). This suggests that PBMC isolated from participants with HIV-2 may produce lower levels of HIV-2 caRNA transcripts following treatment with PKC activators like bryostatin.

Given what we know about differences in the levels of both cell-associated RNA and plasma viremia in ART-naïve individuals with HIV-2, the composition of the HIV-2 enhancer region, and previously published results from SIV-infected macaques, we hypothesized that HIV-2 latency reversal may be more limited than HIV-1 (61). Following *ex vivo* reactivation, we anticipated that PBMC from participants with HIV-2 would produce lower levels of both cell-associated and supernatant RNA compared to HIV-1 PBMC. To test this, we treated isolated PBMC with a selection of latency reversal agents (LRAs) and doses and investigated how HIV-1- and HIV-2-infected PBMCs differed in their LRA responses. We found that, compared to PBMC isolated from HIV-1-infected study participants, HIV-2 PBMC produced significantly lower levels of virion-associated supernatant RNA following treatment with bryostatin, romidepsin, PMA/ionomycin and αCD3/αCD28 beads. Interestingly, a corresponding difference in HIV-2 caRNA levels after LRA treatment was not observed.

Materials & Methods

Study population and sample collection

Samples were obtained from participants enrolled in the HIV Eradication and Latency (HEAL) cohort, a Boston-based longitudinal cohort of participants with HIV, or followed in the HIV clinics of Miriam Hospital in Providence. A total of 13 participants with HIV-1 and 11 participants with HIV-2 were recruited and clinical characteristics of the participants are shown in Table 1 from Chapter 2. All participants gave informed consent prior to enrollment, and had a 90-180cc blood draw into EDTA-containing Vacutainer tubes. Whole blood samples were processed within two hours of collection; plasma was separated and PBMCs isolated by Ficoll-Histopaque density gradient centrifugation before being cryopreserved. Smaller blood draw volumes for a number of HIV-2 participants precluded the isolation of pure resting $CD4^+$ T cell populations, resulting in our decision to perform the reactivation experiments using PBMCs.

LRA treatment conditions

The panel of LRAs used was selected as a result of the different mechanisms of action of the LRAs (Table 2). PBMCs were stimulated with latency-reversing agents at the following concentrations and incubation times: 10 nM Bryostatin-1 for 24 hours; 20 nM bryostatin-1 for 24 hours; 10 nM bryostatin-1 for 48 hours; 20 nM romidepsin for 24 hours; 80 nM romidepsin for 24 hours; 100 nM GS-9620 for 24 hours; 1000 nM GS-9620 for 24 hours; 50 ng/mL PMA + 1 uM ionomycin for 24 hours; 50 ng/mL PMA + 1 uM ionomycin for 48 hours; anti-CD3/anti-CD28 beads (1:1 ratio) for 24 hours; anti-CD3/anti-CD28 beads (1:1 ratio) for 48 hours; or with 5 µl of DMSO for 24 hours as the no-drug control. LRA concentrations were chosen based on previous *ex vivo* studies [22-26]. Cryopreserved PBMCs isolated from HIV-1 or HIV-2 infected individuals

were thawed and rested overnight in R10 media with raltegravir (2 μM) and tenofovir (2 μM) to prevent new rounds of infection. PBMCs were then counted and three to five million PBMCs were seeded into each well in a volume of 5 mL RPMI that includes 10% FBS and 1% penicillin+streptomycin as well as 2 μM raltegravir and 2 μM tenofovir. Each well was treated with a different LRA. Following 24- or 48-hour incubations, cells were harvested, and spun at 1500 rpm for 10 minutes. The supernatant was removed, spun at 2500 for 10 minutes, passed through a 0.22 μm filter, and stored at -80 ºC. Cells were resuspended in 1 mL of RPMI, spun at 1500 rpm for 10 minutes, after which the supernatant was removed via aspiration, leaving only the pellet. Immediately following this, cell-associated RNA and DNA were isolated using AllPrep DNA/RNA Mini Kit (Qiagen).

HIV-2 DNA and RNA qPCR Assay Design

HIV-2 subtype A accounts for the majority of HIV-2 infections. We used the subtype A HIV-2 prototype strain, ROD (GenBank: X05291.1), as our qPCR standard template (106). We selected an HIV-2 region that spans part of the LTR and the *gag* gene for our standard template, and this fragment using primers ROD9475 (CCCGCTAGCTTGCATTGTACTTCG, ROD: 9475-9498) and OG106 (GCCTTCTGAGAGTGCCTGAAATCC; ROD: 1062-1085), which are complementary to regions of ROD that are external to the binding sites of the previously published qPCR primers (146). Platinum Taq High Fidelity Polymerase (Invitrogen) was used for all reactions: 94ºC for 2 m; then 40 cycles of 94ºC for 15 s, 60ºC for 30s, 68ºC for 1 m 30 s; then 68ºC for 5 m. The 1,108-bp amplicon was gel extracted (QIAquick Gel Extraction Kit, QIAGEN) and cloned into the pCR4-TOPO vector TA vector (Invitrogen); the plasmid was electroporated into One Shot® TOP10 Electrocomp *E. coli* (ThermoFisher Scientific) and expanded. DNA

standard templates were generated by amplifying the cloned ROD standard template using M13 primers (FOR: GTAAAACGACGGCCAG; REV: CAGGAAACAGCTATGAC), gel extracted, quantified and serially log_{10}-diluted to generate DNA standards from 10^6 copies to 1 copy. RNA standard templates were generated by linearizing the plasmid with PmeI (New England Biolabs) digestion and transcribing template region RNAs (MEGAscript T7 Transcription Kit, Invitrogen). RNA transcripts were extracted and purified using phenol:chloroform and isopropanol as described, quantified, and serially diluted in DEPC-treated water from 10^6 to 1, including a 5- and 2-copy standard.

Measurements of intracellular HIV

Total DNA and cell-associated RNA were extracted from the centrifuged pellet using AllPrep DNA/RNA Mini Kit (Qiagen) and levels of cell-associated RNA quantified same-day (i.e. no freeze-that cycle) by validated real-time quantitative PCR (qPCR) assays that used a one-step RT-PCR kit (TaqMan Fast Virus 1-Step Master Mix, Applied Biosystems). Total DNA levels were measured by qPCR using TaqMan Universal PCR Master Mix (Applied Biosystems). HIV-1 and HIV-2 persistence markers were quantified using previously published primers and probes (146,147). For HIV-2, we used the qPCR assays described above. Total DNA and cell-associated RNA quantifications are presented as copies of DNA or cell-associated per 10^6 million PBMC. PBMC numbers was determined by measuring CCR5 as previously described (147).

Measurements of supernatant and blood plasma viremia

For RNA extraction of filtered culture supernatant collected from each LRA condition and blood plasma (4.5 mLs for each patient), plasma and culture supernatant were thawed and then

Table 1. LRA Panel

Name	Concentration	Mode of Action	Reference
Bryostatin	10 and 20 nM	Protein kinase C (PKC) agonist; induces reactivation through the NF-kB and AP-1 signaling pathways	(134,142)
Romidepsin	20 and 80 nM	Histone deacetylase (HDAC) inhibitor; increase acetylation of histone proteins promoting transcription	(127)
GS-9620	100 and 1000 nM	TLR7 agonist; promotes secretion of type I interferons (IFNs) that stimulates interferon stimulating gene (ISG) expression	(84,148)
PMA/ionomycin	PMA: 50 ng/mL Ionomycin: 1 μM	PMA: PKC activator; activates T cells inducing cytokine release Ionomycin: calcium ionophore and PKC activator	(144,149),(150,151)
αCD3/αCD28 beads	1:1 bead to cell ratio	T cell activation	(145)

spun at 3000 rpm for 25 minutes at 4 ºC to pellet debris. Plasma supernatants were then transferred to 2 mL Screw Cap Micro Tube (Sarstedt) for ultracentrifugation. Virions were then pelleted by ultracentrifugation at 17,000 rpm for 90 min at 4 ºC. RNA was subsequently isolated from the pellet (QIAamp Viral RNA Isolation Kit, Qiagen). Isolated RNA was immediately quantified using a one-step RT-PCR kit (TaqMan Fast Virus 1-Step Master Mix, Applied Biosystems).

Results

Validation of HIV-2 qPCR Assay

To determine the operating characteristics of our HIV-2 RNA and DNA qPCR assays, we first assessed their specificity and sensitivity. To test specificity for HIV-2, we ran both the RNA and DNA assay with a fragment of ROD spanning the LTR and Gag, serially diluted from 10^6 copies down to 5 copies. Both assays are specific for HIV-2 and did not detect HIV-1 RNA or DNA. The limit of detection of the HIV-2 RNA assay is 10 copies with 100% sensitivity; we also observed 66% sensitivity at 5 copies. For the HIV-2 DNA assay, the limit of detection was also 10 copies with 100% sensitivity; at 5 copies the sensitivity falls to 33%. For HIV-2 RNA, a standard curve generated with triplicate measurements demonstrated an r^2 of 0.9934. For HIV-2 DNA, a standard curve generated with triplicate measurements demonstrated an r^2 of 0.9961.

To assess the reproducibility of the HIV-2 DNA and RNA assays, we performed triplicate qPCR reactions in multiple experiments using identical input HIV-2 RNA and DNA isolated from a representative participant. For HIV-2 RNA, the mean intra-assay coefficient of variation was 2.4% (varied from 0.5 to 4.3%), and the inter-assay variation was 2.8%. For HIV-2 DNA, the mean intra-assay coefficient of variation was 1.1% (varied from 0.9 to 1.3%), and the inter-assay variation was 1.1%.

Fold Increase in HIV-1 and HIV-2 RNA Following Reactivation

To explore how HIV-2 proviruses reactivate from latency and to compare this behavior to HIV-1, we isolated PBMCs from participants infected with either HIV-1 or HIV-2 and stimulated them *in vitro* in the presence or absence of our LRA panel (Table 1). Cell-associated RNA and DNA were extracted from cell pellets, and virion-associated RNA was extracted from culture supernatants; HIV persistence measurements were quantified by qPCR.

We first compared the fold-increase in cell-associated RNA levels after PBMC isolated from HIV-1 and HIV-2 participants were treated with our panel with LRAs (Figure 3.1). Differences in the fold increase of cell-associated RNA between HIV-1 and HIV-2 PBMC were significant only for two conditions: 10 nM Bryostatin for 24 hours ($p=0.007$) and 100 nM GS-9620 for 24 hours ($p=0.03$), where there was a median fold increase of 3.1 and 1.9, respectively, for HIV-2, but only 1.1 and 1.2, respectively, for HIV-1. For all other LRA treatment conditions, there was no significant difference in the fold-increase in cell-associated RNA between HIV-1 and HIV-2 PBMCs. For the remaining nine conditions, median fold increases in HIV-2 cell-associated RNA levels were 1.8 for 20 nM Bryostatin; 1.9 for 10 nM Bryostatin at 48 hours; 2.6 for 20 nM Romidepsin; 3.1 for 80 nM Romidepsin; 1.6 for 1000 nM GS-9620; 6.0 for PMA/ionomycin for 24 hours; 4.6 for PMA/ionomycin at 48 hours; 4.1 for αCD3/αCD28 beads at 24 hours; and 12.4 for αCD3/αCD28 beads at 48 hours. For HIV-1, the median fold increases in cell-associated RNA levels for these nine conditions were 1.0 for 20 nM Bryostatin; 1.2 for 10 nM Bryostatin at 48 hours; 4.6 for 20 nM Romidepsin; 5.2 for 80 nM Romidepsin; 1.4 for 1000 nM GS-9620; 3.7 for PMA/ionomycin for 24 hours; 3.1 for PMA/ionomycin at 48 hours; 2.1 for αCD3/αCD28 beads at 24 hours; and 5.9 for αCD3/αCD28 beads at 48 hours. For the majority of LRA conditions, HIV-1 and HIV-2 PBMCs produced similar levels of cell-associated RNA fold increases (romidepsin

at both 20 nM and 80 nM; PMA/ionomycin at 24 hours and 48 hours). For other LRA conditions, such as αCD3/αCD28 beads at 48 hours, there was a higher fold increase in cell-associated RNA in HIV-2-infected PBMCs.

We then explored if similar patterns were observed at the supernatant RNA level (Figure 3.2). Measurements of supernatant RNA were undetectable or below the limit of quantification for many of the conditions, more so for HIV-2 than HIV-1: 71% versus 37%. We observed that all of the LRA conditions produce comparable fold increases in both HIV-1 and HIV-2 PBMCs, with the exception of romidepsin at 20 nM for 24 hours, where the level of HIV-1 supernatant RNA was higher ($p=0.03$), driven by the large number of undetectable or below the limit of quantification HIV-2 supernatant RNA measurements (only HIV-2 participant 69 had levels of supernatant RN above the limit of quantification).

PBMCs from Participants not on ART Had Similar Levels of RNA post Reactivation

Of 11 HIV-2 participants, three were not on ART at the time of therapy, with two of these participants being treatment naïve. When comparing HIV-1 and HIV-2 reactivation from PBMCs, we only used participants who were virologically suppressed, but given the recruitment of these three participants we assessed whether cell-associated RNA levels could be augmented by LRAs during untreated HIV infection. We treated PBMCs isolated from the three untreated participants with the same panel of LRAs and measured both cell-associated and supernatant RNA levels. All LRA conditions led to average fold increases of 1.5 or higher, with the exception of GS-9620, which led to no increase in cell-associated RNA levels with the exception of one participant (61) where we observed an 8.7-fold increase in cell-associated HIV-2 RNA. Similarly to PBMCs from

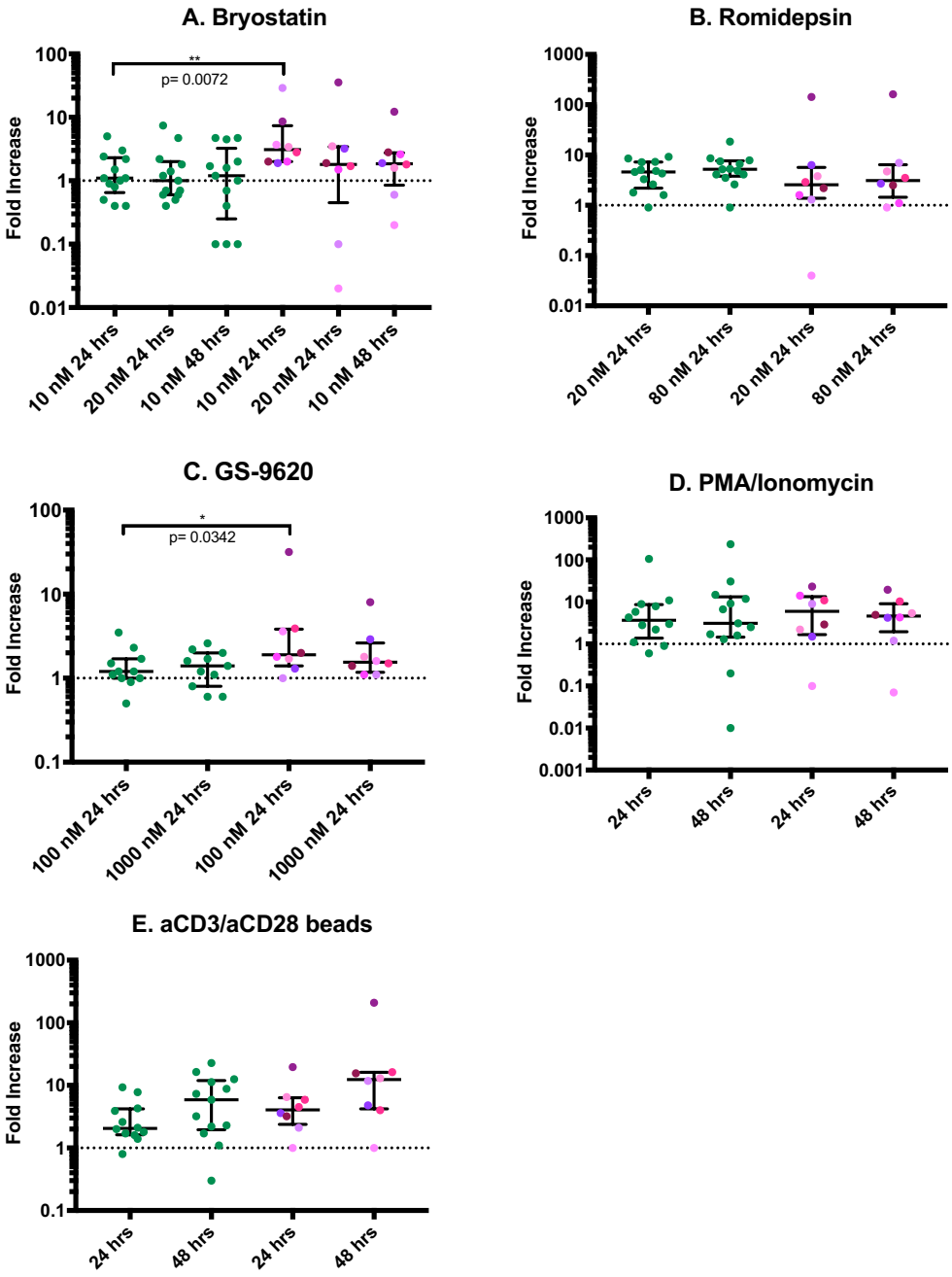

Figure 3.0.1. Fold increase in cell-associated RNA per 10^6 cells compared to no-drug control. Fold increase for HIV-1 (green) and HIV-2 (each color represents a different participant) following reactivation with Bryostatin (A), Romidepsin (B), GS-9620 (C), PMA/ionomycin (D) and αCD3/αCD28 beads (E). The dotted line marks 1-fold increase. Data are presented as a scatter plot with the geometric mean and the geometric SD. * $p \leq 0.05$.

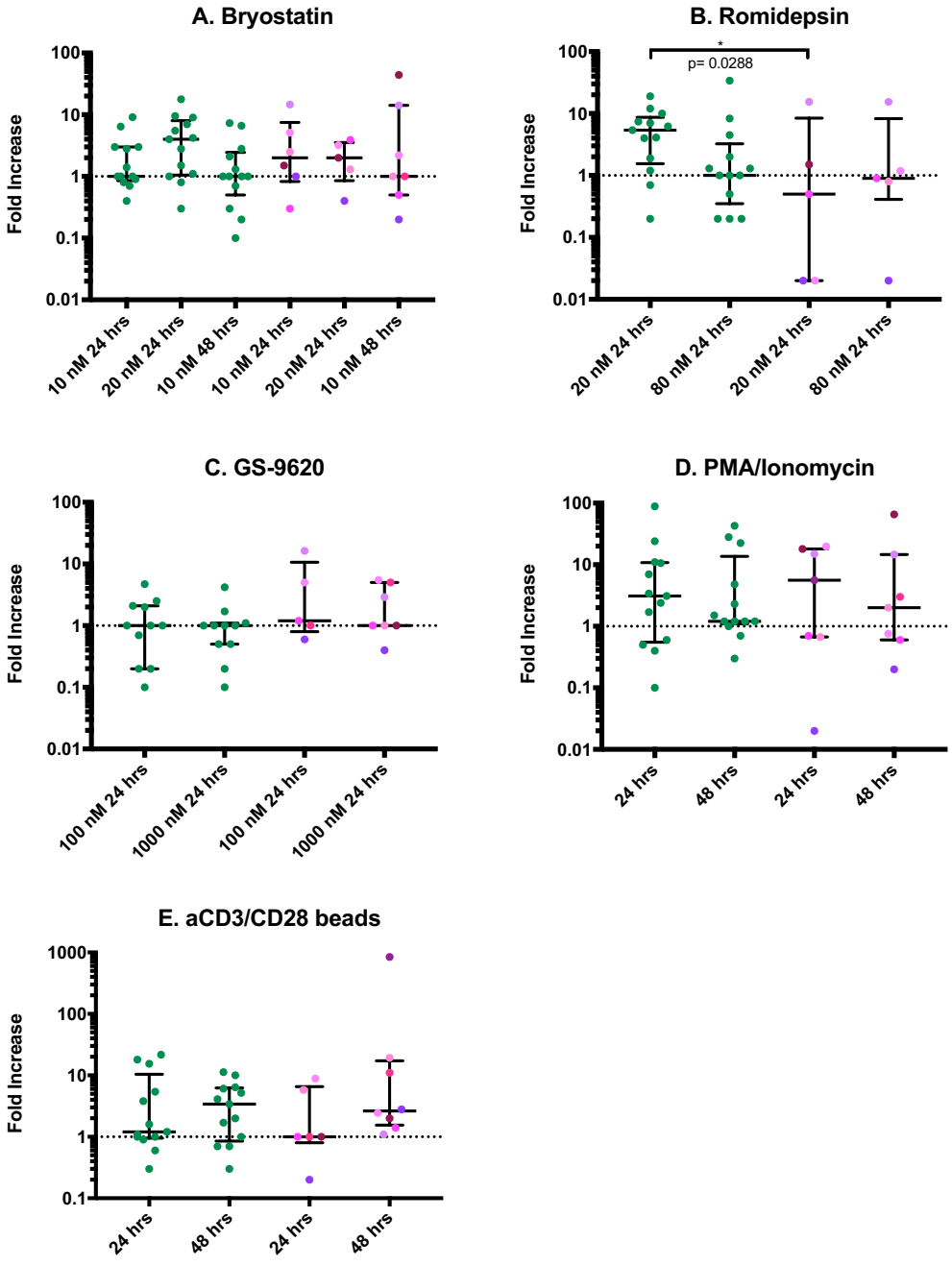

Figure 3.0.2.Fold increase in supernatant RNA per mL compared to no-drug control. Fold increase for HIV-1 (green) and HIV-2 (each color represents a different participant) following reactivation with Bryostatin (A), Romidepsin (B), GS-9620 (C), PMA/ionomycin (D) and αCD3/αCD28 beads (E). The dotted line marks 1-fold increase. Data are presented as a scatter plot with the geometric mean and the geometric SD.

treated HIV-2 participants, measurements of supernatant RNA were undetectable or below the limit of quantification for many of the LRA conditions.

We then compared the fold-increase in cell-associated and supernatant RNA levels between participants on and off treatment (Figure 3.3 and 3.4). At the cell-associated RNA level, there were no observed differences between PBMCs from participants on ART or off ART, with the exception of GS-9620 at 1000 nM, where PBMCs from participants on ART had a significantly higher fold change in cell-associated RNA ($p=0.01$). For other LRA conditions, such as Bryostatin at 10 nM and αCD3/αCD28 beads at 24 hours, we observed a trend where PBMCs from untreated participants yielded lower fold increases in cell-associated production compared to the participants on ART.

The fold increase in levels of supernatant RNA following reactivation of HIV-2-infected PBMCs did not differ significantly between participants on ART and not on ART, nor was there a trend for higher supernatant RNA production in either the treated or untreated PBMCs (Figure 3.4).

Comparing Reactivation in Fresh and Cryopreserved PBMCs

Given the limited number of PBMCs isolated for a number of HIV-2 study participants and the potential for post-thaw cell death, we set up latency reactivation experiments using fresh cells. At the same time, many of the HIV-1 reactivation experiments were done on previously cryopreserved PBMCs. To compare whether fresh and cryopreserved PBMCs behave similarly when treated with our panel of PBMCs, we compared levels of cell-associated RNA (caRNA) per 10^6 million cells following reactivation (Figure 3.5). Except for Heal 58, this comparison was done on PBMCs isolated at different study visits, between 6 months and 1 year apart. While some natural

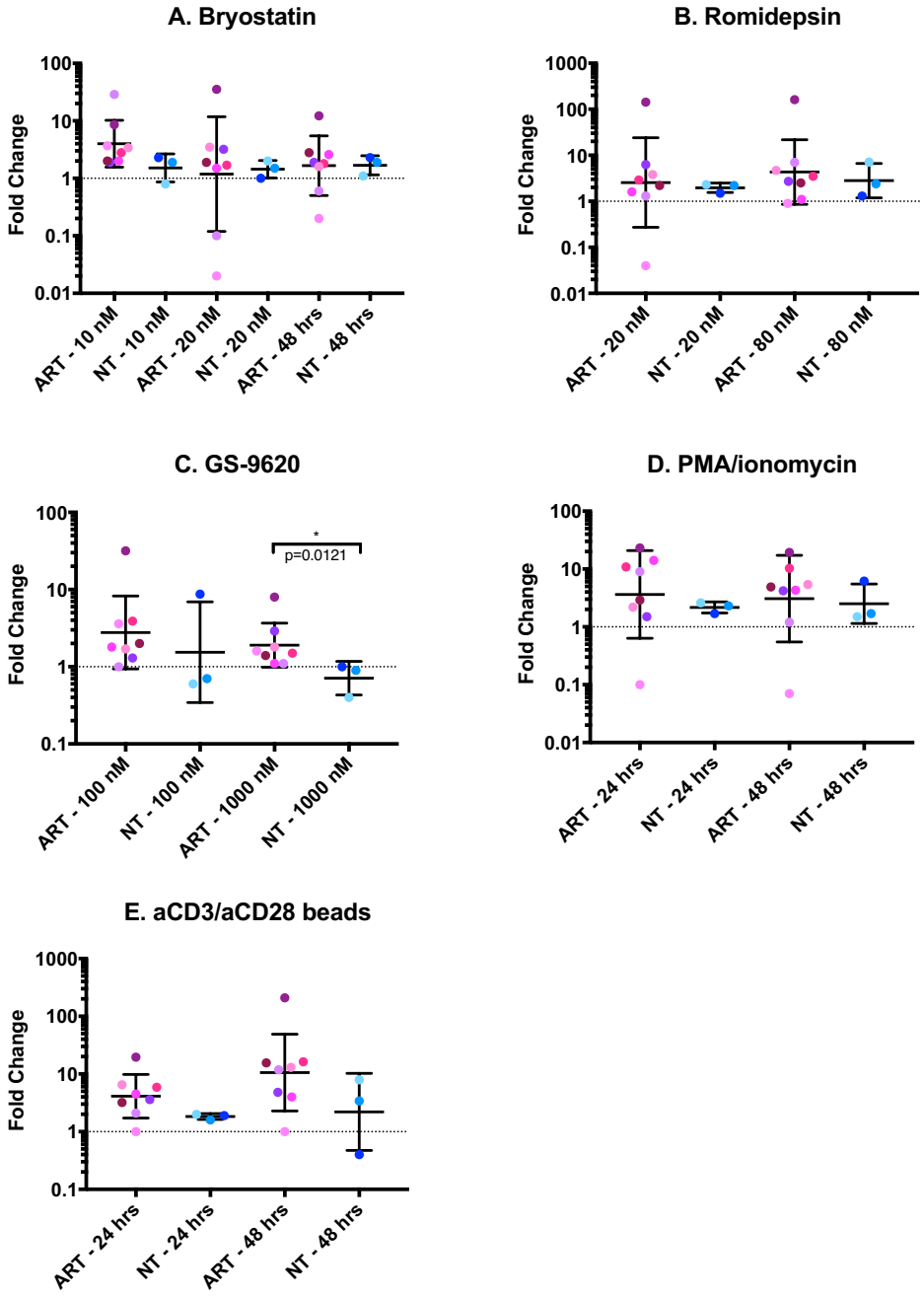

Figure 3.0.3. Fold increase in cell-associated RNA per 10^6 cells. Fold increase compared to no-drug control for HIV-2 PBMCs isolated from participants on ART (pink and purple, each color is specific to a particular participant) and not on ART (blue, each color represents a specific participant) following reactivation with Bryostatin (A), Romidepsin (B), GS-9620 (C), PMA/ionomycin (D) and αCD3/αCD28 beads (E). The dotted line marks 1-fold increase. Data are presented as a scatter plot with the geometric mean and the geometric SD.

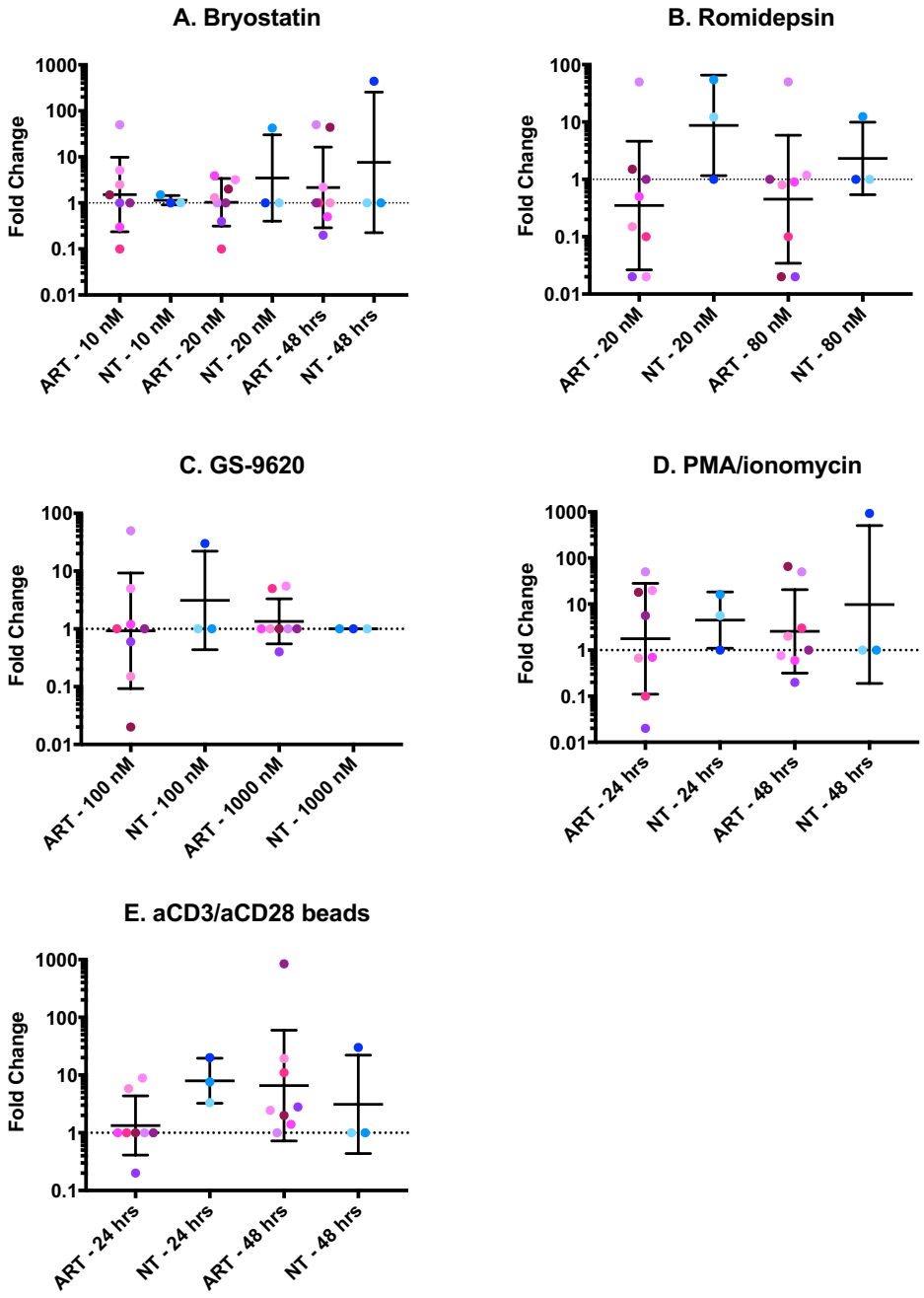

Figure 3.0.4. Fold increase in supernatant RNA per 10^6 cells. Fold increase compared to a no-drug control for HIV-2 PBMCs isolated from participants on ART (pink and purple, each color represents a different participant) and not on ART (blue, each color represents a different participant) following reactivation with Bryostatin (A), Romidepsin (B), GS-9620 (C), PMA/ionomycin (D) and αCD3/αCD28 beads (E). The dotted line marks 1-fold increase. Data are presented as a scatter plot with the geometric mean and the geometric SD.

variation occurs between reactivation replicates, we did not find any significant differences between fresh and frozen PBMCs in terms of cell-associated levels following reactivation. Despite the lack of difference in reactivation when using fresh or frozen PBMCs, we chose to compare reactivation results using cryopreserved and thawed participant PBMCs, with the exception of Heal 57.

Latency Reversal: Normalizing by DNA

While, baseline measurements of genomic DNA for our HIV-1 and HIV-2 study participants did not differ significantly, there was a pattern of higher total DNA levels in the HIV-2 study participants. Therefore, total HIV DNA was measured for each LRA condition to investigate if this pattern persisted in our reactivation experiments. Mean levels of total HIV DNA per LRA condition were significantly higher for HIV-2 PBMCs (1513 ± 84 copies per 10^6 cells, n=8) higher than that measured for HIV-1 participants (127 ± 8 copies per 10^6 cells, n=13; $p<0.0001$). Given that participants were not able to be matched for duration of infection, treatment duration, or CD4 counts, we normalized HIV cell-associated RNA and supernatant viremia measurements for each participant to their total HIV DNA levels.

When normalized by DNA levels, the cell-associated RNA measurements we observed did not differ significantly between PBMCs isolated from HIV-1 and HIV-2 participants following LRA reactivation for all of the LRA conditions tested (Figure 3.6). Normalizing by DNA input also illuminated which LRA conditions were the most effective at producing RNA transcripts following incubation: Romidepsin, an HDAC inhibitor that has consistently been amongst the most efficient LRAs, for both HIV-1 and HIV-2, and αCD3/αCD28 beads for HIV-2.

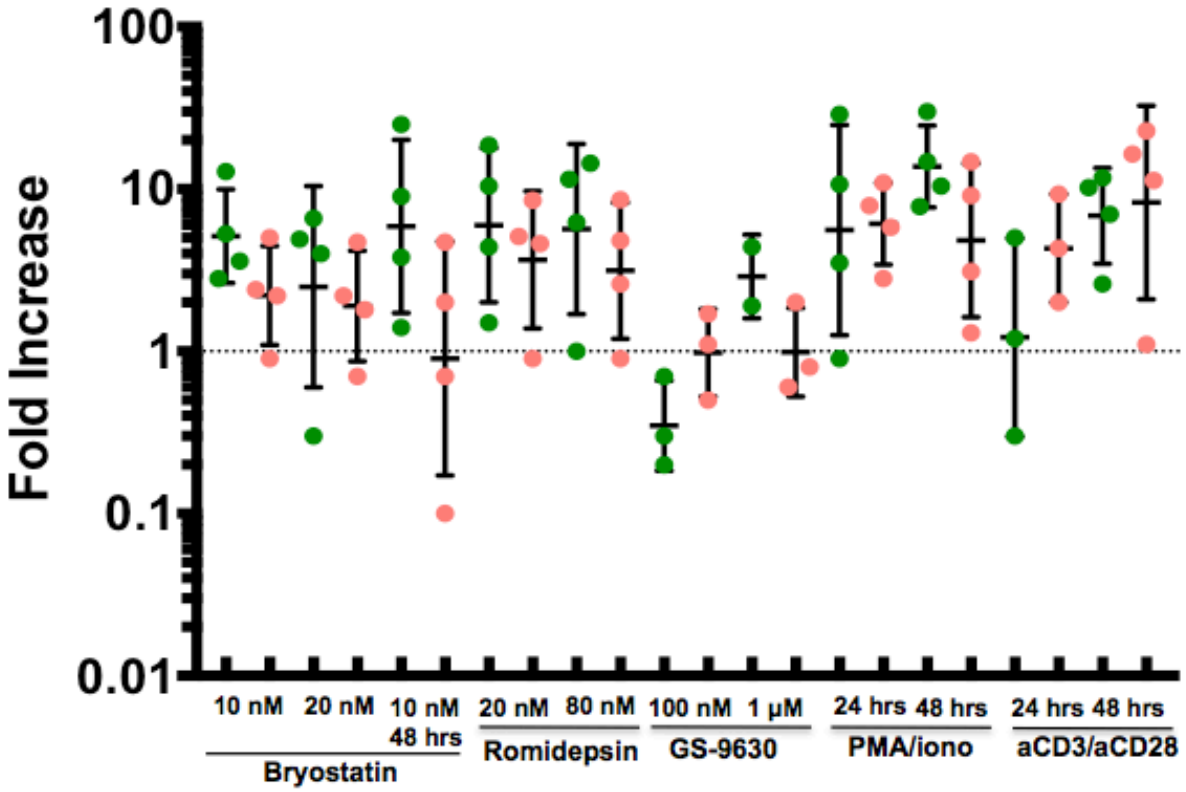

Figure 3.0.5. Fold increase treated versus untreated PBMCs. Comparison of copies of cell-associated RNA per 10^6 cells following reactivation with a panel of LRAs using fresh (green) or cryopreserved-thawed (salmon) PBMCs for four HIV-1 participants: Heal 19, Heal 24, Heal 44 and Heal 58. The geometric mean with the geometric SD is shown for each condition. The dotted line marks 1-fold increase.

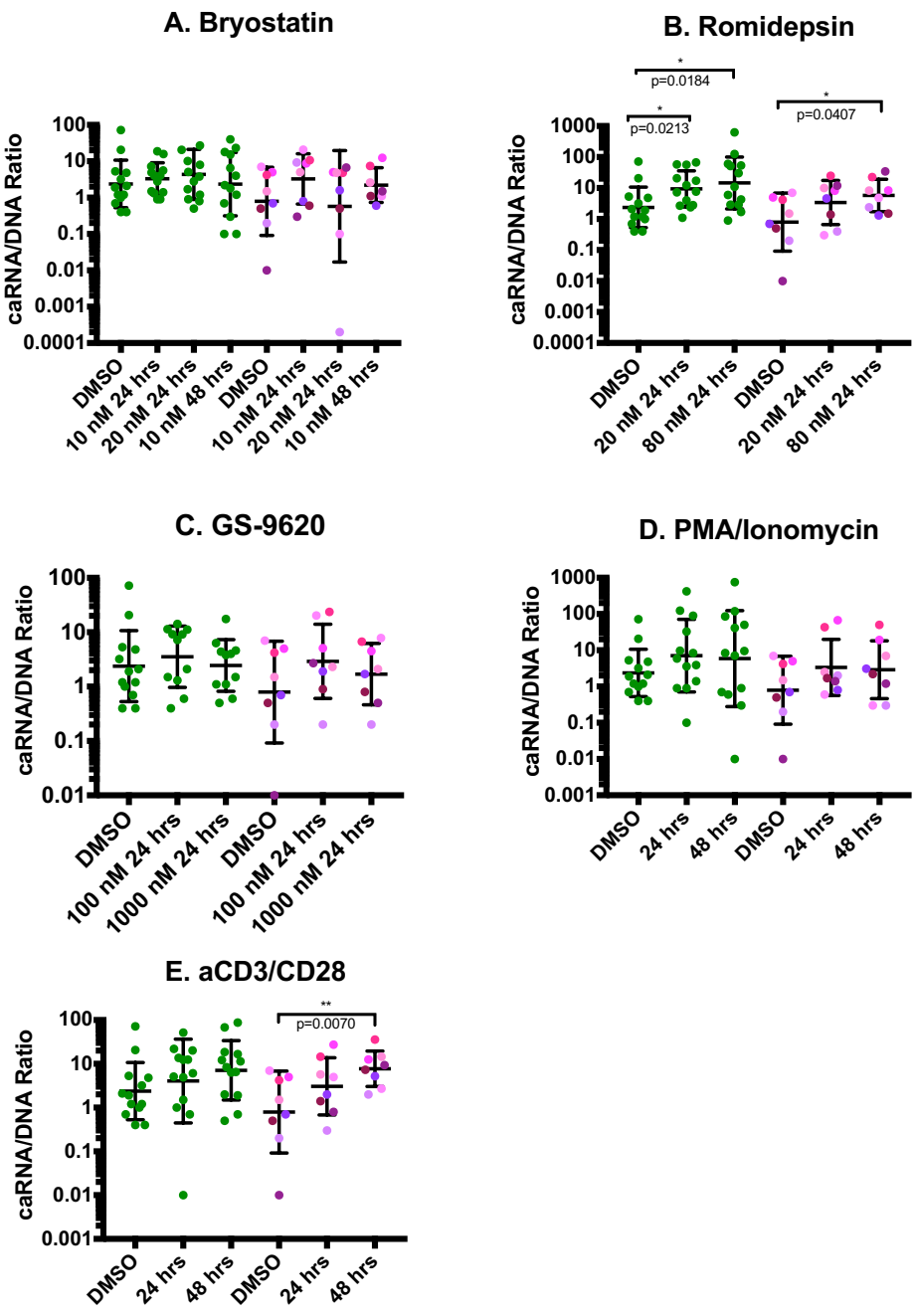

Figure 3.0.6. Cell-associated RNA normalized by DNA. Levels of cell-associated RNA per 10^6 cells normalized by copies of DNA per 10^6 cells for HIV-1 (green) and HIV-2 (each color represents a participant) for study participants following incubation with (A) Bryostatin, Romidepsin (B), GS-9620 (C), PMA/ionomycin (D) and αCD3/αCD28 beads (E). Data are presented as a scatter plot with the geometric mean and SD.

While the levels of caRNA/DNA did not differ significantly between HIV-1 and HIV-2 PBMCs following reactivation for the LRA conditions tested, the levels of HIV-2 supernatant RNA corrected for DNA input, were significantly lower when compared to HIV-1 for specific LRA conditions we tested (Figure 3.7). For Bryostatin at 10 and 20 nM, Romidepsin at 20 and 80 nM, PMA/ionomycin at 48 hours and αCD3/αCD28 beads at both 24 and 48 hours, HIV-1-infected PBMC produced more supernatant RNA than PBMCs isolated from HIV-2-infected participants, when corrected for total HIV DNA levels. Among the remaining LRA conditions where no significant differences between supernatant RNA levels, supernatant RNA/levels following reactivation were higher for HIV-1 PBMCs treated with GS-9620 at 100 nM and PMA/ionomycin at 24 hours. There was no significant increase in HIV-2 supernatant RNA compared to the no-drug (DMSO) control for any of the LRA conditions. For HIV-1 PBMCs, on the other hand, Bryostatin at 10 nM for 24 hours, Romidepsin at 80 nM for 24 hours and PMA/ionomycin for 48 hours, there was a significant increase in HIV-1 supernatant RNA compared to the DMSO control.

Discussion

In a series of *ex vivo* HIV latency reversal experiments, we found that while PBMCs isolated from participants with HIV-1 and HIV-2 that are virologically suppressed can produce cell-associated RNA following reactivation, the magnitude of cell-associated RNA production was higher for HIV-1 PBMCs. Furthermore, HIV-1 and HIV-2 PBMCs produce significantly different levels of supernatant RNA for seven of the 11 LRA conditions tested, with reactivation of HIV-1 PBMCs producing higher levels of virus, while HIV-2 PBMCs often had levels of supernatant RNA that were undetectable or below the limit of quantification.

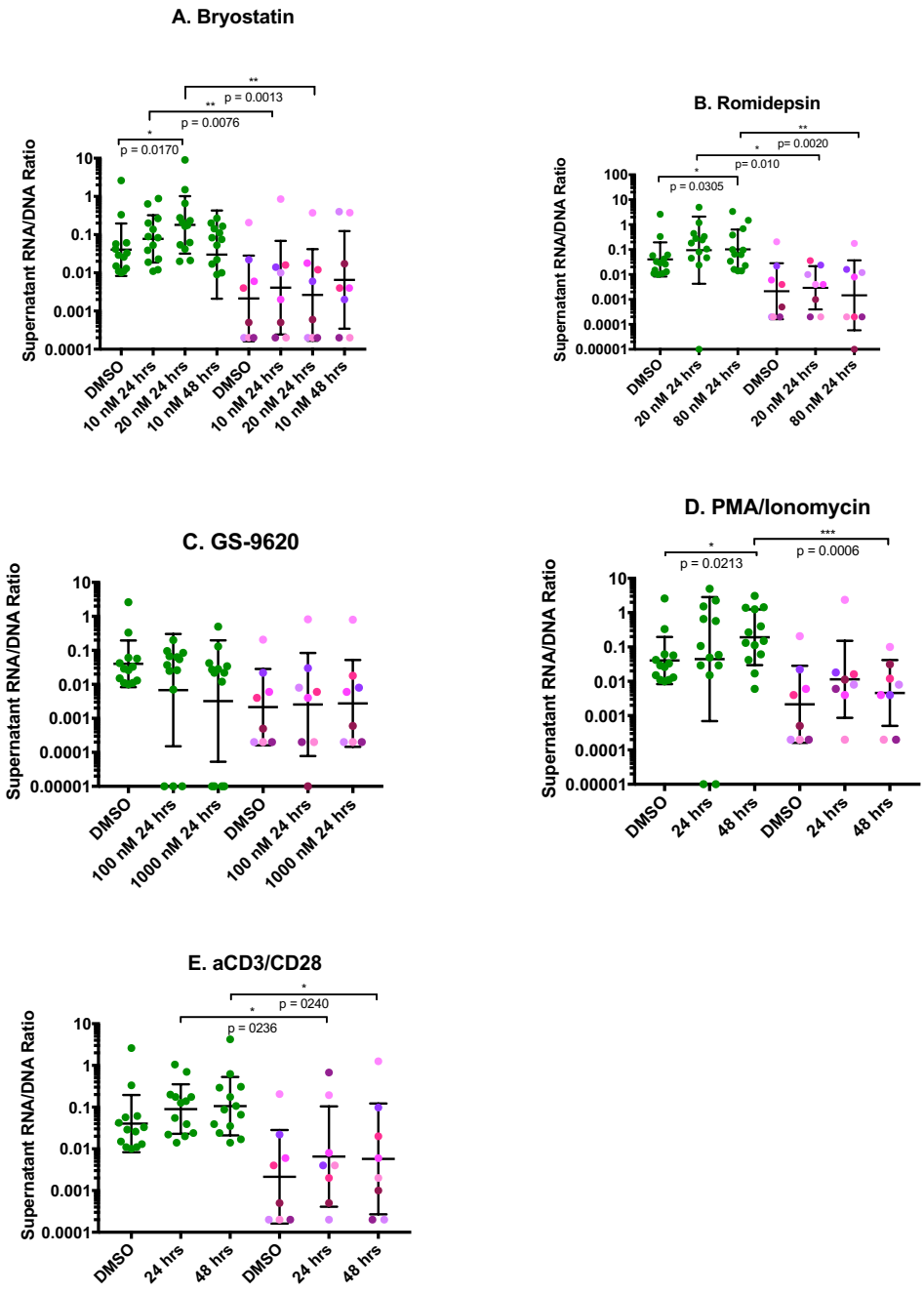

Figure 3.7. Supernatant RNA normalized by DNA. Levels of supernatant RNA per mL normalized by copies of DNA per 10^6 cells for HIV-1 (green) and HIV-2 (each color represents a participant) for study participants. Data are presented as a scatter plot with the mean and SEM. * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$.

Previous measurements of cell-associated RNA and plasma viremia in HIV-1 and HIV-2 study participants have revealed statistically significant differences in RNA levels between these two viruses, while total DNA levels have been comparable (12,117). Baseline measurements of infection for our cohort of HIV-1 and HIV-2 participants on ART show that total DNA levels were not significantly different between the two groups, though HIV-2 participants had a higher level of total HIV DNA per 10^6 cells. At the same time, cell-associated RNA and plasma viremia levels are not significantly different between HIV-1 and HIV-2 infected participants, probably as a result of these participants being on ART. Despite the lack of significant differences in either cell-associated RNA or plasma viremia levels between our HIV-1 and HIV-2 study participants on treatment, we anticipated that reactivated HIV-1 and HIV-2 PBMCs would produce levels of cell-associated RNA and supernatant RNA that mirror what has been observed in ART-naïve study participants. While we observed a statistically significant difference between measurements of HIV-1 and HIV-2 supernatant RNA, levels of HIV-1 and HIV-2 cell-associated RNA were not statistically different, a result that was different from our hypothesis. This became more pronounced when cell-associated RNA and supernatant RNA measurements were normalized by total DNA.

Given that the rarity of virologically suppressed HIV-2 participants on ART precluded our ability to match our cohort to HIV-1 participants with similar $CD4^+$ T cell counts, normalizing by DNA allowed us to examine the production of RNA transcripts and virions while controlling for DNA input. When we do not control for DNA input, fold increase data suggests that HIV-2 PBMCs are equally efficient at producing RNA transcripts and virions following reactivation with our panel of LRAs. Additionally, the fold increase in cell-associated RNA is actually higher (and in the case of Bryostatin at 10 nM and GS-9620 at 100 nM, significantly so) in HIV-2 PBMCs than

HIV-1 PBMCs for the majority of LRAs with the exception of Romidepsin. By normalizing by total HIV DNA, the levels of which have been shown to correlate with the number of replication-competent virus in *ex vivo* viral outgrowth, we get a better sense of how LRAs affect latent HIV-1 and HIV-2 PBMCs (152).

We observed comparable levels of HIV-1 and HIV-2 cell-associated RNA normalized by DNA for all LRA conditions, even Bryostatin. As a PKC activator, Bryostatin reverses latency through the NF-κB and AP-1 signaling pathways, and HIV-1 has more Bryostatin-responsive transcriptional elements in its enhancer region. The HIV-2 enhancer, on the other hand, lacks AP-1 binding sites and contains only a single functional NF-κB site. Despite this, we observed similar levels of caRNA/DNA following treatment with all Bryostatin conditions. This may be because the concentrations of Bryostatin used (10 nM and 20 nM) were elevated enough to make up for the difference in the enhancer regions of HIV-1 and HIV-2. It is possible that at lower Bryostatin concentrations, a difference emerges in the ability of PBMCs isolated from HIV-1- and HIV-2-infected study participants to produce cell-associated RNA from latency.

HIV-1 and HIV-2 PBMCs treated with Romidepsin, PMA/ionomycin and αCD3/αCD28 beads behaved similarly to those treated with Bryostatin: no difference at the caRNA/DNA level, while HIV-1 produced higher levels of supernatant RNA/DNA. The TLR7 agonist GS-9620 did not produce significantly different levels of HIV-1 or HIV-2 at either the cell-associated RNA or supernatant RNA level, when these were normalized by DNA.

Through our cohort, we had access to three HIV-2 participants who were not on ART: one was briefly on ART during pregnancy, while the other two were ART naïve. We compared the ability to reactivate these HIV-2 infected participants to our latently infected HIV-2 participants and found no differences in their levels of cell-associated or supernatant RNA following

reactivation with our panel of LRAs. This may be because these three participants, despite not being on ART, had undetectable plasma RNA levels at the time of blood draw. This perhaps suggests some amount of virologic control, by an unknown mechanism, independent of ART. Undetectable plasma RNA is a common characteristic among individuals with HIV-2 and a large percentage of those infected with this virus behave as long-term non-progressors (115). A smaller percentage of individuals infected with HIV-1 are considered long-term non-progressors or elite controllers; as such, it's unclear whether a similar behavior would be observed in PBMCs isolated from infected HIV-1 participants, or if this is limited to participants who have achieved some level of virologic control.

There are limitations to this study, both in terms of the size of the study cohort and the lack of matching between HIV-1 and HIV-2 study participants. For one, the HIV-2 participants had significantly higher levels of total DNA compared to the HIV-1-infected participants; this difference may be the result of a longer seeding period, reflecting an extended time between HIV-2 infection and the participant being placed on ART. Furthermore, there are no clinical guidelines for treating people living with HIV-2, and many participants go through a number of ART regimens before arriving at their current one, thus further extending the reservoir seeding period. This scenario is particularly likely for participant Heal 32, who had total DNA level of 4285 copies per 10^6 cells, as well as Miriam 104, with a measured 9208 copies per 10^6 cells. At the same time, the limited number of HIV-2 participants precluded our ability to match HIV-1 and HIV-2 participants in terms of $CD4^+$ T cell count or other factors such as age, duration of infection, and sex. A study with more HIV-2 participants on ART matched to HIV-1 participants may find no significant difference in total DNA levels between HIV-1 and HIV-2 participants. Similar

situations have occurred where the levels of total DNA were compared between unmatched ART naïve HIV-1 and HIV-2 participants (153,154).

Work done in ART-naïve study participants by Kanki and colleagues observed growing differences in the levels of HIV-1 and HIV-2 genomic material as they went from total DNA to cell-associated RNA and finally plasma viral loads (117). Specifically, as the integrated provirus is transcribed and eventually translated to make new virions that are released from the infected cell, levels of HIV-1 become increasingly higher compared to HIV-2. Our reactivation data shows the same thing, with the exception of the total DNA levels which we've addressed earlier in the discussion. One of the more surprising patterns seen in our results is the disconnect between RNA transcripts and virion production observed in reactivated HIV-2 PBMCs. Our DNA normalized results show that both HIV-1 and HIV-2 latently infected PBMCs are capable of producing similar levels of RNA transcripts following reactivation. At the same time, we also find a marked difference in their ability to produce supernatant RNA, with many HIV-2 PBMCs producing supernatant RNA post-treatment that was undetectable or below the limit of quantification. This suggests a potential post-transcriptional block that affects the ability of reactivated HIV-2 PBMCs to transform RNA transcripts into virions; Kanki and colleagues postulated the possibility of a post-integration block, and our data suggests that such a block would be acting at either a post-transcriptional or post-translational step. This block, for instance, could occur at translation, assembly or budding. Further research, particularly work investigating the exact nature of the RNA transcripts produced following reactivation, will begin to parse out where such a block may be acting. Furthermore, should such a block be found, it may point to new tools for preventing HIV-1 transcription in a "block and lock" HIV control strategy.

CHAPTER 4

HIV-1 proviral landscape characterization varies by pipeline analysis

HIV-1 proviral landscape characterization varies by pipeline analysis

Fernanda A. Ferreira[1,2], Qianjing He[2], Stephanie Banning[2], Olivia Roberts-Sano[2], Olivia Wilkins[2], Daniel R. Kuritzkes[2,3], Athe M. N. Tsibris[2,3§]

1 Virology Program, Graduate School of Arts and Sciences, Harvard University, Cambridge, MA, USA
2 Division of Infectious Diseases, Brigham and Woman's Hospital, Boston, MA, USA
3 Harvard Medical School, Boston, MA, USA

Keywords: proviral landscape, reservoir, HIV-1, characterization

Abstract

Introduction: HIV persists in $CD4^+$ T cells and rebounds after ART cessation, representing a barrier to cure. To better understand the residual proviral DNA reservoir, research groups have designed experimental approaches that identify provirus sequences as intact or defective; analysis pipelines further differentiate the nature of defective proviruses. We investigated the effects that different analysis pipelines had on the identification and characterization of HIV-1 proviral sequences.

Methods: We used single genome amplification (SGA) to amplify near full-length HIV-1 DNA sequences isolated from peripheral blood mononuclear cells (PBMC) of thirteen treated, suppressed participants with HIV-1. Amplicons were directly sequenced by next-generation single genome sequencing (NG-SGS) and analyzed using four HIV-1 proviral characterization pipelines.

Results and discussion: HIV-1 sequences derived from thirteen participants were studied. A total of 693 HIV-1 proviral sequences were generated, 283 of which were near-full length. We found that proviruses often harbor multiple sequence defects and that the elimination order used by each

pipeline affected the percentage of proviruses identified and allotted into each defect category. These differences varied across participants, depending on the number of defect categories to which a given provirus sequence could be assigned.

Conclusions: The choice of pipeline used for HIV-1 proviral DNA landscape analysis may bias the characterization of defective sequences. To compare landscape characterization results across research groups and studies more easily, the development of a consensus elimination pipeline should be considered.

Introduction

During HIV-1 infection, viral DNA integrates into the genomes of $CD4^+$ T cells and can persist during antiretroviral therapy (97). Previous work has investigated what proportion of integrated sequences are intact and therefore may contribute to the replication-competent HIV-1 reservoir (85,88,89,98,107). The vast majority, greater than 90%, of integrated proviruses are defective, accumulate during acute HIV infection, and cannot support virus replication (85,89). When characterizing intact proviral sequences, the nature of the defective proviruses – whether due to large deletions, hypermutation, inversions, premature stop codons or other categories of defects – is typically also reported. These various types of defects arise through different mechanisms that include HIV reverse transcriptase-induced mutations, template switching during reverse transcription, and APOBEC3-mediated cytidine deamination that causes guanine-to-adenine (G-to-A) hypermutation (85,155,156).

Current proviral landscape research focuses primarily on the dichotomy between defective versus intact sequences; defective proviruses are typically further characterized into distinct defect categories. These "defective" proviruses may still retain transcriptional and translation activity that

is relevant to HIV eradication efforts (88,113). In general, research groups that characterize provirus sequences use custom in-house analysis pipelines. Some labs categorize amplified DNA sequences manually, whereas others have developed automated, scripted pipelines (98,107). While most pipelines include all the classic provirus defects (e.g. hypermutation, large deletions, etc) the order in which these defective proviral forms are eliminated differs. It is therefore possible that different research groups will arrive at similar conclusions as to the numbers of intact proviruses but the proportion of proviruses assigned into each defect category may vary. If true, direct comparisons of HIV provirus characterizations across research groups may be difficult and could compromise a deeper, meta-analytic understanding of the defective nature of the HIV reservoir. Furthermore, recent work has moved beyond just intact and defective, focusing on the forces that shape the proviral landscape such as cytotoxic T lymphocytes (CTLs), which may preferentially target cells with hypermutated or PSI/MSD defective proviruses (113).

To determine the effect of analysis pipeline on defective HIV-1 proviral characterization, we categorized the proviral landscape of 13 HIV-1 participants using published elimination pipelines from different research groups and explored how the proportion of sequences assigned to each defect category varied as a function of the particular analysis pipeline that was used. We found that proviral sequences can typically be assigned into multiple defect categories. Not only does the proportion of sequences within each defect category change with the pipeline used but that this may have an unexpected effect on the reported percentage of intact HIV-1 proviral sequences.

Methods

Participant Cohort

Cryopreserved peripheral blood mononuclear cells (PBMC) were obtained from thirteen participants with HIV-1 enrolled in the HIV Eradication and Latency (HEAL) cohort, a Boston-based longitudinal cohort of participants with HIV. Peripheral venipuncture samples were collected between 2015-2018 and PBMC were isolated by Ficoll-Hypaque density gradient centrifugation. All participants gave informed consent prior to enrollment; the HEAL biorepository study was approved by the Partners Human Research Committee.

Proviral Amplification, Sequencing and Analysis

Total DNA isolated from participant PBMC samples was subjected to limiting dilution prior to near-full-genome amplification with Platinum Taq HiFi polymerase (Thermo Fisher Scientific). HIV-1 proviruses were amplified using previously published PCR conditions and primers (98,105). The primers were: first-round forward primer: 5'-AAATCTCTAGCAGTGGCGCCCGAACAG-3' (HXB2: 623-649); first-round reverse primer: 5'-TGAGGGATCTCTAGTTACCAGAGTC-3' (HXB2: 9662-9686); second-round forward primer: 5'-GCGCCCGAACAGGGACYTGAAARCGAAAG-3' (HXB2: 638-666); and second-round reverse primer: 5'-GCACTCAAGGCAAGCTTTATTGAGGCTTA-3' (HXB2: 9604-9632). A convenience endpoint of 20 amplified near full-length sequences (>8000 bp) per participant was set. This goal was attained for all participants with the exception of HEAL 09, where low levels of total HIV-1 DNA limited the number of near-full-length amplicons to nine.

PCR amplicons were directly sequenced by next-generation single-genome sequencing (NG-SGS) using an Illumina deep sequencing platform. Sequenced proviruses were manually

analyzed and classified as either containing a large deletion (<8000 bp), hypermutated, carrying a 6 nucleotide or longer deletion in the 5'-untranslated region (UTR) deletion, carrying an internal deletion (INDEL), having inversions, having a stop codon deletion, carrying a frameshift mutation, carrying a deletion in one of the four stem loops of the packaging signal (PSI) and/or a mutation in the multiple splice donor (MSD) site, or intact. Only sequences with length greater than 8,000 bp were analyzed for the other defect categories. Briefly, proviruses with APOBEC3G-mediated hypermutation were identified using the Los Alamos National Laboratory HIV Sequence Database Hypermut algorithm (108). The Los Alamos National Laboratory HIV Sequence Database Gene Cutter tool was used to identify stop codons and frameshift mutations (157). INDELs were identified using the Los Alamos National Laboratory HIV Sequence Database HIVAlign tool and defined as an equal or greater than 2% difference in the sequence of each HIV gene (109). 5'-UTR deletions (set at 6 or more deletions), PSI deletions and MSD mutations were identified by aligning all the sequences to HXB2 (Geneious Prime 2019.1.1) and checking for deletions and mutations (110). Inversions were identified by aligning each proviral sequence to HXB2 and comparing them using the Dotplot viewer (110). Proviral sequences that lacked any of these defects were labeled as intact.

Results and Discussion

Participant Characteristics

For the thirteen participants we studied, the mean CD4 count was 807 cells/mm^3 (SD=393). Participants were receiving antiretroviral therapy for at least 6 months and all had plasma HIV-1 RNA levels that were below the limit of detection of a commercially available virus load assay. Participant characteristics are shown in Table 1.

Table 1. Characteristics of Study Subjects

Patient ID no.	Sex	Age	$CD4^+$ T Cell count ($cells/mm^3$)	Therapeutic Regimen*	Months with Viral Load < 50 copies/mL
01	M	54	606	3TC, ABC, EFV	24
04	M	37	773	DRV/cobi, FTC TAF	>80
09	F	32	1713	EVG/cobi, FTC, TDF	44
16	M	56	267	ETR, FTC, RAL, TDF	10
19	M	54	416	DRV/r, DTG, FTC, MVC, TAF	6
24	M	55	691	EVG/cobi, FTC, TAF	12
25	M	31	984	EVG/cobi, FTC, TAF	12
30	M	61	1005	EFV, FTC, TDF	>27
38	M	52	695	DTG, FTC, TAF	>27
44	M	58	448	DTG, FTC, TAF	16
50	M	65	899	DTG, FTC, RPV, TAF	42
57	M	58	1346	EVG/cobi, FTC, TAF	17
58	M	45	647	DRV/cobi, FTC, TAF	29

*3TC, lamivudine; ABC, abacavir; BIC, bictegravir; DRV/cobi, cobicistat boosted darunavir; DRV/r, ritonavir boosted darunavir; DTG, dolutegravir; EFV, efavirenz; ETR, etravirine; EVG/cobi, cobicistat boosted elvitegravir; FTC, emtricitabine; MVC, maraviroc; RAL, raltegravir; RPV, rilpivirine; TAF, tenofovir alafenamide; TDF, tenofovir disoproxil fumarate

Characterization of HIV-1 Proviral Defects

We assessed the proviral characterization pipelines from three research groups, as well as our own, based on their published descriptions and personal communication (85,98,107). We used the same definition of a given defect category across all pipelines, unless noted, but acknowledge that our stringency may differ from that originally used by each research group. A comparison of the four pipelines is shown in Figure 4.1. Whereas all pipelines generate the number of intact proviral sequences in the final sequence step, there is marked heterogeneity in preceding steps. Large deletions and hypermutated sequences are generally identified earlier in the processes but their order of characterization varies. Inversions are categorized either first, in the penultimate step, or are not explicitly categorized. Defects in the 5'-untranslated region (UTR) may be handled as deletions anywhere along its length or more specifically as deletions in the PSI packaging signal or mutations at the major splice donor (MSD) site. An evaluation for frameshift mutations, nonsense mutations, stop codons and insertions/deletions (indels) occur in all pathways but that is where the similarities end. The order of these steps varies, as do the permutations on how they may be performed in one combined sequence step.

To understand how different analysis pipelines affect sequence identification and categorization, we used each of the pipelines to analyze 693 HIV-1 proviral sequences isolated from thirteen HEAL cohort participants. An average of 53 sequences were analyzed per participant; 12 of 13 participants had at least 20 near full-length sequences. Of the 283 near-full-length proviral sequences, 230 (81.3%) carried more than one type of defect and the distribution of the number of defect categories per provirus sequence varied across participants (Figure 4.2). On average, a defective sequence could be assigned to 2.7 (95% CI 2.5, 3.0) defect categories. For some participants, e.g. HEAL 04, 09 and 58, all defective sequences were assigned to two or more

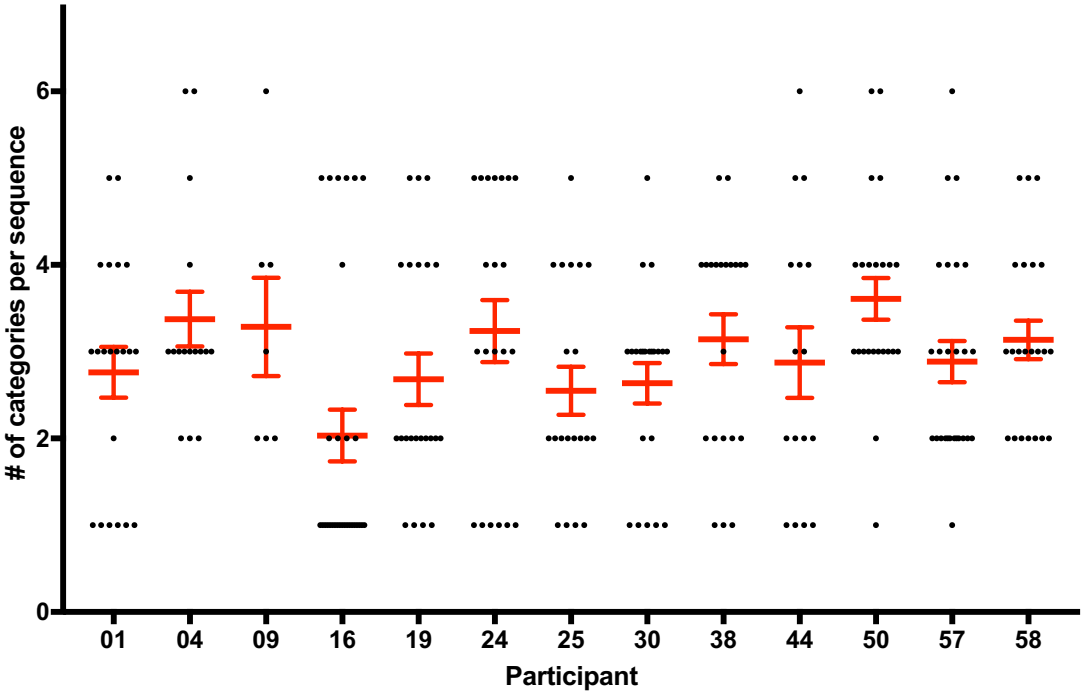

Figure 4.0.1. Comparison of the number of categories assigned to each amplified sequence for every study participant. For amplified provirus sequences >8,000 bp there are six defect categories they can be assigned to: hypermutation, 5'-UTR deletion, PSI deletion and/or MSD mutation, INDEL, inversion and stop codon and/or frameshift. Each point represents a single amplified sequence and the mean and SEM are shown in red.

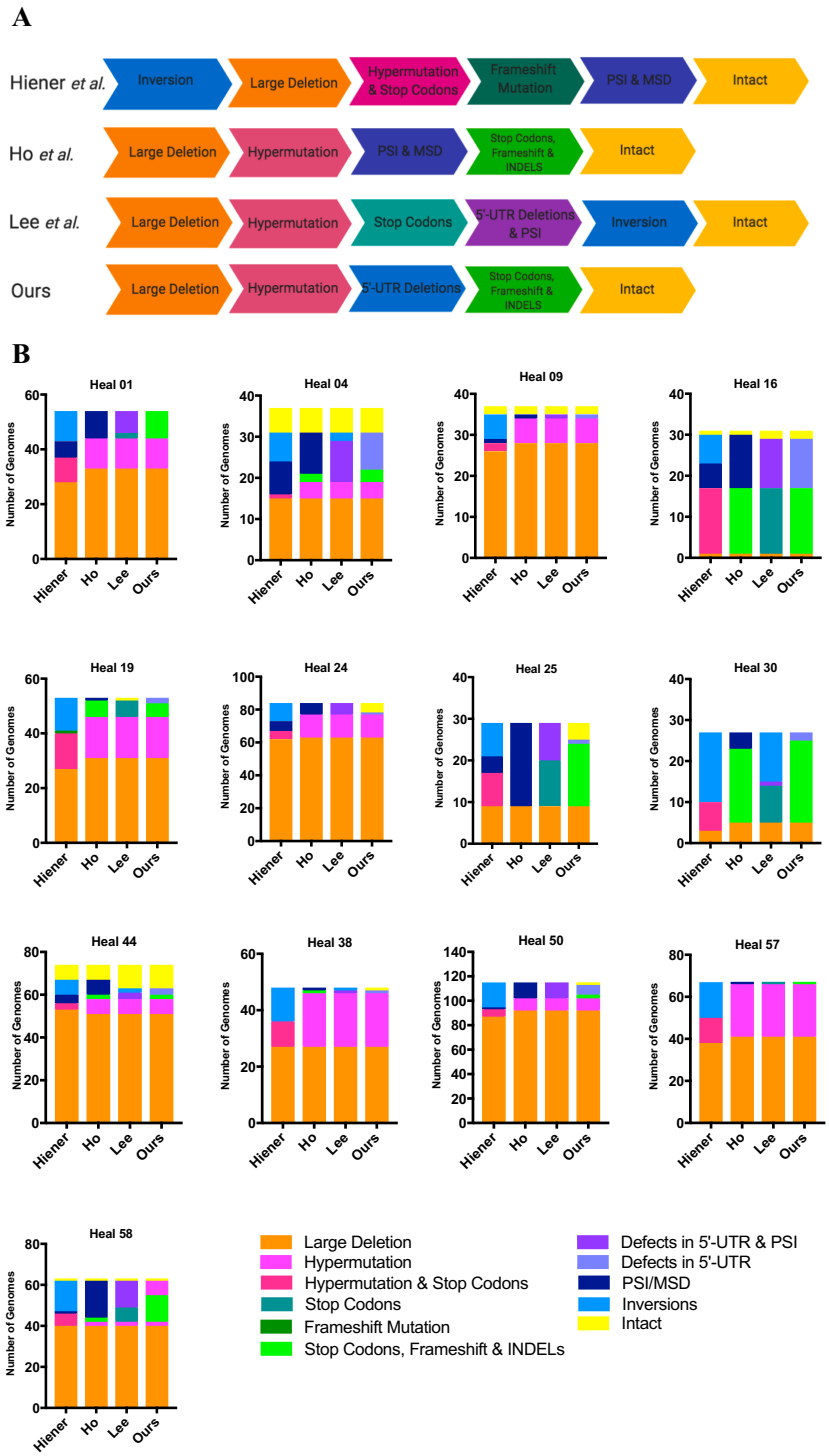

Figure 4.0.2. Characterization pipelines and their effect on the proviral landscape of HIV-1-infected participants. **A**. Comparison of the different processes of elimination used by various research groups to identify intact, potentially replication-competent HIV-1 proviruses. Adapted from Ref. 1, 2 and 8. **B**. Number of genomes allotted into each category within the total pool of sequenced proviruses by each pipeline investigated. The four pipelines are, from left to right, Hiener et al., 2018; Lee et al., 2017; Ho et al., 2013; and our own pipeline.

defect categories. For HEAL 16, on the other hand, 63% of defective near-full-length sequences were assigned to a single category, with the remaining sequences assigned to two, four or five categories. Following the assignment of defects to each proviral sequence, we compared the proviral landscape of all 13 participants generated with each pipeline (Figure 4.2).

We observed an association between certain defect categories. Hypermutation leads to stop codons, which is why some pipelines combine the two categories or characterize only hypermutation (98,107). However, for certain participants (e.g., HEALs 01, 16, 25, 30 and 58) we observed instances where stop codons occurred without hypermutation. In HEALs 16, 25 and 30, for example, none of the near-full-length sequences with stop codons displayed G-to-A hypermutation. For HEAL 01 and 58, 2 and 4 near-full-length sequences, respectively, had stop codons not related to hypermutation. By looking only at hypermutation, sequences with stop codons that were not the result of APOBEC-mediated G-to-A mutations may be considered intact or placed in a different defect category. Alternatively, combining hypermutation and stop codons into a single category of defects may inflate the percentage of the proviral landscape for a participant that is deemed hypermutated.

We also observed an association between 5'-UTR deletions, PSI deletions and MSD mutations. Amongst our 13 study participants, there were only two instances where there is a perfect overlap between sequences containing 6 or more deletions in the 5'-UTR and those containing a deletion in the PSI stem loops and/or a mutation at the MSD site: HEALs 01 and 38, where one near-full-length proviral sequence falls into this general defect category in each pipeline investigated. For the other 11 participants, this is not the case; here we observe that the more specific the category – using a combined PSI deletion and MSD mutation instead of the general 5'-UTR deletion category – the higher the number of sequences that are considered defective in

this general 5'-UTR category. For instance, for HEAL 01, the number of defective sequences is zero when looking at just 5'-UTR deletions, eight when combining 5'-UTR and PSI deletions, and ten when combining PSI deletions with MSD mutations. This is also the situation with HEAL 25, where defective near-full-length sequences go from one when looking globally at the 5'-UTR deletion to 20 when assessing for both PSI deletions and MSD mutations. Even just adjusting this category to include deletions in the PSI, already increases the number of sequences in this category for HEALs 01 (0 to 8), 04 (9 to 10), 24 (1 to 7), 25 (1 to 9), 50 (8 to 13), and 58 (0 to 13). In the case of certain pipelines, looking only at 5'-UTR deletions leads our pipeline to consider near-full-length sequences intact, while the other pipelines categorize them as defective. The non-inclusion of MSD mutations causes pipelines to erroneously inflate the number of intact proviruses in HEALs 16 (from 1 to 2 intact), 24 (0 to 6 intact), 25 (0 to 4), 38 (0 to 1 intact), 44 (7 to 11 intact), and 50 (0 to 2 intact). This demonstrates that the specific pipeline used can influence not just the number of provirus sequences in each category, but also the number of sequences considered intact.

The Effect of Pipeline on Defective Sequence Proportions

Given that 81.3% (230 sequences) of proviral sequences over 8000 bp contain two or more types of defect, the order of elimination used by a pipeline, as well as the inclusion or exclusion of certain categories, may influence the final proviral landscape. To evaluate this, we calculated, on a per-participant basis, how the percentage of sequences assigned to each category changes as a function of the pipeline used (Figure 4.2). We also looked at the effect of the pipeline on the cumulative proviral landscape for all 13 participants (Figure 4.3).

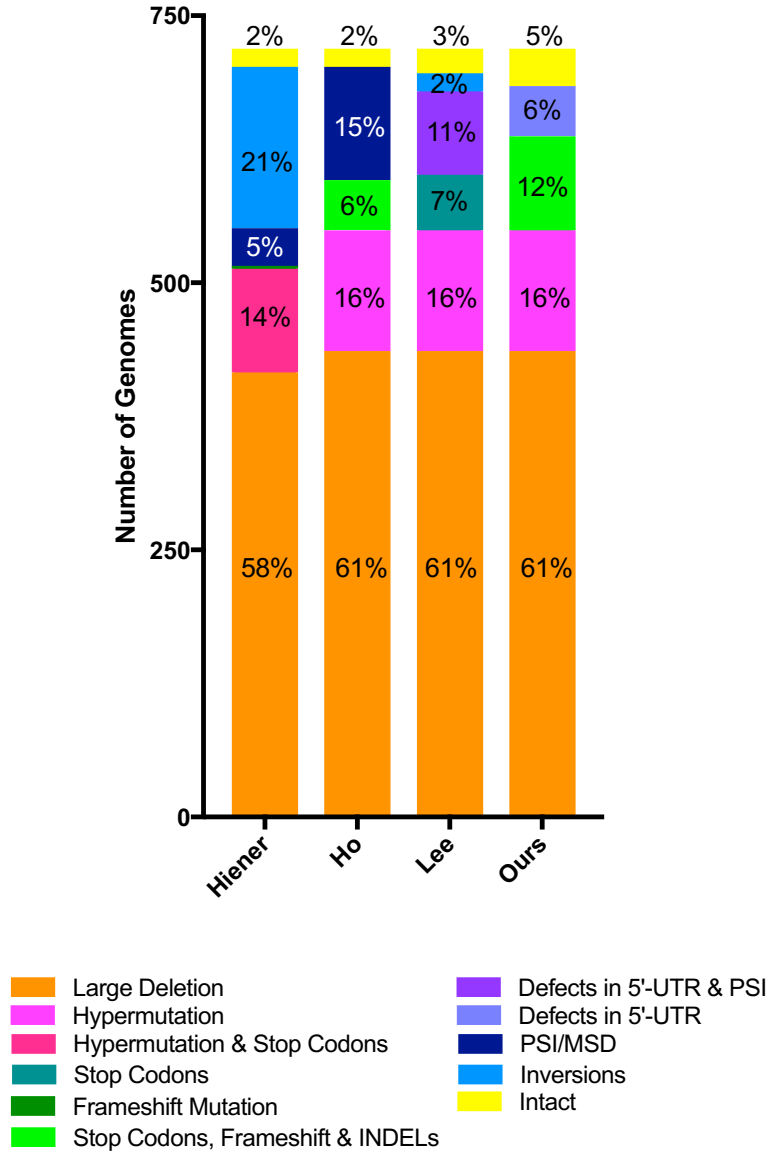

Figure 4.0.3. Comparison of the cumulative landscape for the study cohort generated using the four pipelines. Number of genomes in each defect category for all 693 proviral amplicons according to the four pipelines analyzed, showing the percentage of genomes in each category.

Certain pipelines combine HIV-1 sequence defect categories and this has an influence on the final landscape. “Stop Codons” are their own category in Lee *et al.*, but are combined with “Hypermutation” in Hiener *et al.* and with “Frameshift Mutations” and “INDELs” in Ho *et al.* HEAL 16 is the only participant where this has no effect on the number of genomes in this category: there are always 16 sequences assigned independently of whether “Stop Codons” are combined with any other type of defect. HEAL 16 is unique in this respect because 63.3% of HEAL 16 proviruses have a single defect. For the other 12 participants, however, the combination of “Stop Codons” with other defects leads to slight discrepancies in the percentage of sequences in each category, as exemplified by HEAL 38 where the number of sequences in a “Stop Codon” category varies from 9 to 0 depending on the pipeline.

The elimination order of the pipeline also influences the final percentage of sequences in each category and this is particularly evident in the case of “Hypermutation.” All pipelines examine for the presence of hypermutated sequences, though the order in which it is assessed varies. The further down in the pipeline the “Hypermutation” category is, the fewer sequences are categorized as hypermutated. For instance, for HEAL 38, 19 provirus sequences are hypermutated according to three of the pipelines, but this number falls to 9 when sequences are checked for inversions before hypermutations. The same happens for HEALs 01 (11 to 9), 04 (4 to 1), 09 (6 to 2), 19 (15 to 13), 24 (14 to 5), 44 (7 to 3), 50 (10 to 6) and 57 (25 to 12).

An “Inversions” defect category is included in two pipelines, though the order in which it appears is different in each: inversions are the first defect category in the Hiener *et al.*, pipeline, but the last category in Lee et al. Unsurprisingly, while the Hiener *et al.* pipeline leads to 21% of all genomes categorized as having inversions when we look at the cumulative of all 13 participants, this percentage falls to 2.4% when using the Lee *et al.* pipeline.

Given the different pipelines and that the majority of near-full-length genomes carry at least two types of defects, it's not surprising that the cumulative proviral landscape varies according to the pipeline used (Figure 4.3). Across all pipelines, the "Large Deletion" category has similar total number of genomes: 436 for all pipelines, with the exception of Hiener et al. where there are 416 genomes. The percentage of intact and hypermutated genomes is also similar across the four pipelines (between 2.4% and 4.9% for intact and 13.5 and 15.7% for hypermutated) and in line with what has been previously shown for HIV-1 (85,98,107,113). For the other categories, however, the number of genomes in each category differ substantially, the result of how specific a category is (e.g., PSI deletion/MSD mutations vs. 5'-UTR deletions) and at what stage in the pipeline it is assessed. Looking generally at defects in the 5'-UTR, the number of genomes in this region varies from 37 (5.1%) to 106 (14.8%) depending on the specificity of the defect category and when it appears in the pipeline. The same is seen with defect categories involving "Stop Codons," which can represent anywhere from 6.5 to 12.2% of the proviral landscape.

Conclusion

Published strategies for quantifying intact sequences in the HIV-1 reservoir use a step-by-step elimination pipeline that analyzes each sequence amplified to see whether it has a specific defect, with sequences that have no apparent defect classified as intact. A byproduct of these elimination pipelines is a measurement of the ways in which the provirus is defective. As research begins to explore more aspects of the proviral landscape, moving beyond intact versus defective, this byproduct information will prove increasingly useful. Therefore, it is important to know if we can compare percentages of defects across research groups. In this small study we explored

whether the elimination order used influences both the proviral landscape of an individual participant (Figure 4.2) as well as the cumulative proviral landscape of a study cohort (Figure 4.3).

Our results suggest that the elimination order can influence the number of genomes allotted into each defect category, as well as the number of sequences deemed "Intact." This occurs because the majority of amplified provirus sequences are defective in more than one way (Figure 4.1) and, as such, the elimination order used by a pipeline has discrete effects on the number of sequences in each category. The exact categories used also influences the final proviral landscape for each participant characterized and, in general, the more specific a defect category, the fewer sequences that are deemed intact by the end of the pipeline, as seen in the case of 5'-UTR where using PSI deletions and MSD mutations is a more stringent approach. Work exploring how CTLs shape the proviral landscape have looked at PSI/MSD defects; we found that the percentage of genomes in the 5'-UTR defect group varied from 5 to 15% depending on the pipeline, something to be aware of when looking at landscape results from various research groups. .

Our results show that it is important to be aware that the characterization pipeline used has an effect on the proviral landscape generated and suggests that the field would benefit from having a consensus elimination pipeline. Such a pipeline should analyze amplified sequences for all the defect categories, allowing for the identification of provirus sequences that carry multiple types of defects and generating information about how defect categories overlap or not for each study participant. While this approach may lead to results that are more complicated to analyze, it illustrates better the nature of the proviral landscape and would eliminate the ordering issue that leads different elimination pipelines to generate discretely different reservoir landscapes.

CHAPTER 5: DISCUSSION

New Insights into the HIV-2 Reservoir

The work presented here extends our understanding of the HIV-2 reservoir in terms of the nature of the provirus and its dynamics when treated with latency reversing agents (LRAs). Very little work to characterize the proviral HIV-2 landscape has been previously done and our work confirms and expands on recent findings (93). Future work, with larger participant cohorts and access to greater numbers of PBMCs, thus permitting the isolation of $CD4^+$ T cells , will allow us to follow up on the mechanisms responsible for the observations we have made and add additional details. Work done in the HIV-1 reservoir field may provide a "blueprint" in this regard.

The HIV-2 Proviral Landscape: What Determines the Nature of the HIV Reservoir?

Before Bender *at al.* published their characterization of the HIV-2 proviral reservoir of three ART-treated participants with HIV-2 from Senegal, it was unknown whether the HIV-2 proviral landscape would look more like HIV-1, where in general over 90% of the proviruses are defective, or SIV, where approximately 25-30% of the proviruses are intact (93). While total HIV DNA levels have been shown to be comparable between ART-naïve, HIV-1 and HIV-2 study participants, HIV-2 is more closely related to SIV, with which it has about 75% sequence homology (5,117). Our investigation of the proviral landscape of nine HIV-2-infected and 13 HIV-1-infected study participants builds on the work of Bender *et al.* and, despite amplifying the proviral genomes using different methods and characterizing them with slightly modified pipelines, our results are comparable to theirs and together demonstrate that the HIV-2 proviral landscape more closely resembles that of HIV-1-infected individuals (93). Our dataset is composed of 678 proviral sequences, compared to their 41 sequences, allowing us to not only confirm their finding but extend on it as well as increase confidence in the results. The cumulative fraction of intact HIV-2 proviruses for all study participants was 2.4% (1 out of 41 amplicons) for Bender *et*

al. and 8.7% (59 out of 678 amplicons) for our HIV-2 participants (93). The higher percentage of intact amplicons found in our study participants is largely driven by HEAL 31, a distinct participant. Of the 31 near-full-length genomes amplified for this participant, 30 were intact and there is evidence that the reservoir maintenance of HEAL 31 is primarily driven by clonal expansion: 10 amplicons and 11 amplicons form two clusters that differ due to a single mismatch. When HEAL 31 is removed from the cumulative proviral landscape analyses, the fraction of intact genomes falls to 4.5% (29 out of 647), which is closer to what Bender *et al.* observed (93). HEAL 31 also exemplifies the benefits of investigating the nature of the HIV-2 proviral reservoir in a larger set of participants, thus allowing one to capture the full range of reservoir behaviors. In our cohort of nine HIV-2 participants, HEAL 31 is an outlier, but it is possible that if one were to expand the cohort other participants with a mostly intact reservoir would be identified. Beyond the intact versus defective dichotomy, our results also confirm the fraction of amplicons harboring large deletions and hypermutation. Large deletions were present in 66% and 65.3% of viral genomes amplified from the three Senegal participants and our participant cohort, respectively (93). The fraction of proviral genomes harboring G-to-A hypermutations was also similar: 12% versus 11.4% (93).

Our exploration of the proviral landscape of these nine HIV-2-infected participants gives a more complete picture of the HIV-2 reservoir. The nine participants we studied are on a range of treatment regimens, including integrase inhibitors, nucleoside reverse-transcriptase inhibitors (NRTIs), and protease inhibitors, while the three participants in Bender *et al.*, were on a protease inhibitor and nucleoside analog regimen; this suggests that the treatment regimen is not determining the proviral landscape. Furthermore, our participant cohort includes a possible HIV-2 group B participant (HEAL 69), which is the less common of the two HIV-2 groups, and we

found no difference in the proviral landscape of HEAL 69 compared to the other HIV-2 participants, suggesting that the subtype of the virus does not play a role in shaping the proviral landscape. The greater diversity of our cohort and the comparable numbers of intact proviral genomes across the HIV-2 study participants, suggests that, in general, the proviral landscape of HIV-2-infected individuals is composed primarily of defective proviruses independent of treatment regimen or HIV-2 group. Exceptions do exist, however, and we identified one participant – HEAL 31 - with a largely intact, probably clonally expanded HIV-2 provirus landscape.

In our work, we characterized the proviral landscape of thirteen participants with HIV-1 and used them as a comparator group, tailoring our pipeline to make it as similar between HIV-1 and HIV-2, The number of intact HIV-1 genomes we found (17 out of 693, or 2.4%) is in the range that has been previously reported by other groups characterizing the HIV-1 reservoir.(85,88–90,98,99) It is also similar to the proportion of intact proviral genomes reported for three participants with HIV-2 (2.4%). Given how much less common individuals with HIV-2 are in the USA, we could not match the baseline characteristics of our HIV-1 and HIV-2 participant groups. The two groups differed by sex (92.3%, and 45.5% male for HIV-1 and HIV-2, respectively) and race (38.5% black for HIV-1, and 81.8% black for HIV-2). Despite this, we still found a similar fraction of the provirus landscape to be intact in participants with HIV-1 or HIV-2. Considering how distinct HIV-1 and HIV-2 are at a sequence-level (40% nucleotide sequence homology) and their divergent collection of accessory proteins (vpu vs. vpx), our results suggest that host factors, rather than viral factors, play a major role in determining the look of the viral reservoir in humans.

Pollack *et al.* suggested one host factor that may have a role in shaping the proviral landscape (113). Using data from two previously published studies, they showed that the proportion of intact and ψ/MSD defective proviruses did not correlate with the duration of time

between diagnosis and enrollment, meaning that the number of these types of proviruses did not change significantly over time (113). The proportion of proviruses with large internal deletions and hypermutations, however, did change, with the number of hypermutated proviruses going down while the number of proviruses with large internal deletions going up. This data supports the dynamic nature of the HIV provirus, and suggests that the processes involved in determining the proviral landscape may affect each type of provirus defect in a distinct manner.

In the maintenance of the HIV reservoir during treatment, two distinct processes have been the focus of most research groups: the proliferation of latently infected cells, resulting in expanded cells that contain the integrated provirus, and ongoing viral replication despite ART, which sees novel infections of $CD4^+$ T cells (112,158). Clonal expansion, which has been shown in HIV-1-infected patients and for which we found evidence in HIV-2- participants, 31, 69, 84 and 104 can result from the homeostatic proliferation of infected $CD4^+$ T cells (98)(90). This clonal expansion, which can be driven by antigenic stimulation, results in one viral variant dominating the proviral landscape. Ongoing viral replication, on the other hand, is contested and may only make a limited contribution to the maintenance of the reservoir (159).

All of the known processes involved in shaping the HIV proviral landscape – homeostatic proliferation, immune pressure, CTLs – are host factors and, as such, the same processes are present in both HIV-1 and HIV-2 infected individuals. Further research will need to confirm our evidence of clonal expansion in HIV-2-infected participants, and to explore whether defective HIV-2 proviruses can also be expressed and, in turn, if these cells are then recognized by CTLs. Nonetheless, the similarities in the proportion of provirus categories between HIV-1 and HIV-2 suggests that the processes that act on the HIV-1 reservoir also shape the HIV-2 reservoir.

HIV-2 and the potential of a post-integration block

One of the aspects of HIV-2 infection that has garnered considerable research attention is its lower pathogenicity compared to HIV-1. When HIV-2 was first isolated in 1986, HIV-2 patients with AIDS had not yet been observed (2). Even once it was determined that HIV-2 infections could progress to AIDS, it was clear that the rate of disease development is lower following HIV-2 infection (3,18). In HIV infections. pathogenicity correlates with the level of virus in the plasma, and when Popper *et al.* measured and compared the plasma viral load of ART-naïve HIV-1 or HIV-2 sex workers in Senegal, they found that the median viral load was 30 times lower in women infected with HIV-2 (12). MacNeil *et al.* expanded on this work, measuring the total DNA and cell-associated RNA levels, as well as the plasma viral load (117). They found that the levels of cell-associated RNA and plasma viral load were significantly higher in participants infected with HIV-1 compared to those infected with HIV-2 (117). The levels of total DNA, however, were not different between HIV-1 and HIV-2 participants (117). This led them to suggest that the difference in pathogenicity of the two viruses could be due to different replication rates and that HIV-2, which is equally capable to of seeding a reservoir, may have a higher propensity for latency.

Latency is established and maintained primarily through a number of different mechanisms, including transcriptional regulation. Transcriptional regulation involves the absence of host transcription factors such as NF-κB, Sp1, AP-1 and Ets-1 amongst others (160). HIV-2 has a different collection of transcriptional enhancers compared to HIV-1 and Tong-Starksen *et al.* demonstrated that reporter plasmids containing the HIV-2 enhancer region were less responsive to activation signals than plasmids with the HIV-1 region (61). If HIV-2 does have a higher propensity for latency, it is possible that the makeup of the enhancer region plays a role. Another mechanism involved in maintaining latency is chromatin remodeling, with nucleosomes placed

near the transcriptional start site (TSS) when inactive; these nucleosomes are displaced during activation (161). When the integration sites of HIV-1 and HIV-2 were compared, integration into heterochromatin, was more common in HIV-2 participants than HIV-1 participants (57). As such it is also possible that chromatin remodeling, particularly the ability to make the provirus more available to transcription factors following activation, plays a role in regulating the HIV-2 proviruses ability to escape from latency.

In Chapter 3, we investigated, for the first time, the ability of HIV-2 to reactivate from latency using PBMC isolated from participants with HIV-1 or HIV-2 . We used a panel of latency reversing agents (LRAs) across a range of mechanisms of action, including PKC activators, which reverses latency through the NF-κB and AP-1 signaling pathways, and one histone deacetylase inhibitor (HDACi), which leads to chromatin relaxation. Because mean levels of HIV-2 total DNA per LRA condition were significantly higher, we normalized both cell-associated RNA and supernatant RNA levels by DNA. When normalized by DNA, post-reactivation levels of cell-associated RNA did not differ significantly between PBMCs isolated from HIV-1 and HIV-2 participants for any of the LRA conditions tested. This is different from what MacNeil *et al.* found when they measured cell-associated RNA levels in ART-naïve participants with HIV-1 or HIV-2 from Senegal: levels of cell-associated RNA were significantly higher in HIV-1 participants (117). It is also somewhat unexpected given how the enhancer region of HIV-2 has only a single functional NF-κB binding site and no AP-1 binding sites (61). Bryostatin, a PKC activator, signals through both pathways and, yet, when we normalize by DNA, PBMC from HIV-1 and HIV-2 study participants produce comparable levels of cell-associated RNA. Romidepsin, an HDACi, is the only LRA that leads to a level of production of both HIV-1 (at 20 and 80 nm) and HIV-2 (only at 80 nM) cell-associated RNA that is significantly higher than the DMSO control. When

normalized by DNA, HIV-1 and HIV-2 PBMCs produce comparable levels of cell-associated RNA, though HIV-1 PBMCs produce significantly higher levels of supernatant RNA.

At the supernatant level, when RNA measurements are normalized by total DNA, levels of HIV-1 supernatant RNA were significantly higher than HIV-2 for seven of the 11 LRA conditions. These seven conditions include Bryostatin (10 and 20 nM), Romidepsin (20 and 80 nM), PMA/ionomycin (48 hours) and αCD3/αCD28 beads (24 and 48 hours). These conditions include PKC activators (Bryostatin and PMA), an HDAC inhibitor (Romidepsin) and a T cell activator (αCD3/αCD28 beads), showing that despite different mechanisms of action, LRAs lead to a higher production of HIV-1 virions compared to HIV-2 virions. Furthermore, levels of supernatant RNA were only significantly higher than the DMSO control for HIV-1 PBMCs following treatment with Bryostatin, Romidepsin and PMA/Ionomycin. These results more closely resemble what has been observed for ART-naïve, study participants, where the plasma viral load was significantly higher in HIV-1 than in HIV-2 participants.

When MacNeil *et al.* quantified measurements of HIV-2 infection in ART-naïve study participants, they concluded that HIV-2 replication is attenuated after DNA integration, and a potential block affects some aspect of HIV transcription. Our reactivation data, which sees a significant difference between HIV-1 and HIV-2 only at the supernatant RNA stage, places this potential block further down the replication cycle, possibly at translation. *In vitro* work suggests that the translation of HIV-2 mRNAs is less efficient, due in a large part to structural differences at the 5'-UTR between the two viruses; specifically, the trans-activation response element (TAR), which is a single stem-loop in HIV-1, is a larger, more complex structure in HIV-2 that appears to interfere with the viruses ability to effectively translate mRNAs (63). Another possibility is rather

than a single block, both HIV-2 transcription and translation are inefficient compared to HIV-1 in multiple ways that cumulatively determine the lower viremia of HIV-2.

A Consensus Provirus Characterization Pipeline

In Chapter 4 we compared four proviral characterization pipelines that use step-by-step elimination to analyzes each proviral amplicon to see whether it harbors a specific defect, with sequences that have no apparent defect classified as intact. We found that 81.3% of amplicons carry two or more defects and as such the order of elimination can influence the number of genomes allotted into each defect category, as well as the number of sequences that are considered intact. Furthermore, the definition used for each category also influences the number of sequences placed into a defect category. In fact, following the pipeline comparison, we modified our own pipeline for both HIV-1 and HIV-2 proviral characterization, replacing the more inespecific "5'-UTR defect" for "PSI deletion and/or MSD mutation." As such, we suggested that the HIV reservoir field would benefit from establishing a consensus pipeline. While we don't prescribe an exact formulation for the consensus pipeline, we recommend that each defect category be as specific as possible and that each amplicon be analyzed for every defect, thus identifying provirus sequences that carry multiple types of defects.

Concluding Remarks

Our data adds new insights into the HIV reservoir, both in terms of the composition of the characterization of the provirus and its behavior following reactivation. A core question that infuses all HIV-2 research is whether the virus' lower pathogenicity is due to host or viral factors. In Chapter 2, our results confirm work from Bender *et al.*, demonstrating that the proviral

landscapes of HIV-1 and HIV-2 participants have similar proportions of intact proviruses (2% and 9%, respectively), as well as similar proportions of the common defect categories. This similarity suggests that the forces shaping the proviral landscape are host factors, rather than viral factors. Furthermore, the HIV-2 sequence is more homologous to SIV than HIV-1; despite that, the SIV landscape is 28% intact, adding further credence to host factors playing a role in shaping the proviral landscape.

In Chapter 3, we compared levels of cell-associated RNA and supernatant RNA following reactivation of PBMCs isolated from HIV-1 and HIV-2 study participants. Our data, when normalized by DNA, showed that HIV-1 PBMCs saw significantly higher levels of supernatant RNA than HIV-2 PBMCs for six of the reactivation conditions, despite not observing a similar dynamic at the cell-associated RNA level. This suggests the existence of a block that prevents mRNA transcripts from becoming virions; a post-integration block was previously proposed by MacNeil *et al.*, after they measured DNA, cell-associated RNA and plasma viremia in HIV-1 and HIV-2 ART-naïve study participants. Since our experiments were conducted *ex vivo*, without the pressures of host immunity, this block is likely due to viral factors, such as the TAR element of the HIV-2 5'-UTR.

The research presented here points towards two avenues of future research, one focused on host factors and the other on viral factors. On the proviral reservoir end, it will be interesting to confirm whether host factors such as CTLs, which have been suggested to play a role in HIV-1, determine the HIV-2 proviral landscape. In terms of reactivation and the potential block, one can explore where a potential block is occurring and whether it is due to viral factors.

REFERENCES

1. Barin F, Denis F, Allan JS, M'Boup S, Kanki P, Lee TH, et al. Serological Evidence for Virus Related to Simian T-Lymphotropic Retrovirus III in Residents of West Africa. Lancet. 1985;2(8469–70):1387–9.

2. Kanki PJ, Barin F, M'Boup S, Allan JS, Romet-Lemonne JL, Marlink R, et al. New human T-lymphotropic retrovirus related to simian T-lymphotropic virus type III (STLV-IIIAGM). Science (80-). 1986;232(4747):238–43.

3. Clavel F, Guétard D, Brun-Vézinet F, Chamaret S, Rey MA, Santos-Ferreira MO, et al. Isolation of a new human retrovirus from West African patients with AIDS. Science (80-). 1986;233(4761):343–6.

4. Brun-Vezinet F, Katlama C, Roulot D, Lenoble L, Alizon M, Madjar JJ, et al. Lymphadenopathy-Associated Virus Type 2 In AIDS And AIDS-Related Complex: Clinical and Virological Features in Four Patients. Lancet. 1987;1(8252):128–32.

5. Chakrabarti L, Guyader M, Alizon M, Daniel MD, Desrosiers RC, Tiollais P, et al. Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature. 1987;328(6130):543–7.

6. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med. 2011;1(1):a006841.

7. D'arc M, Ayouba A, Esteban A, Learn GH, Boué V, Liegeois F, et al. Origin of the HIV-1 group O epidemic in western lowland gorillas. Proc Natl Acad Sci. 2015;112(11):E1343–52.

8. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Rev Investig Clin. 1983;220(4599):868–71.

9. Santiago ML, Range F, Keele BF, Li Y, Bailes E, Bibollet-Ruche F, et al. Simian Immunodeficiency Virus Infection in Free-Ranging Sooty Mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: Implications for the Origin of Epidemic Human Immunodeficiency Virus Type 2. J Virol. 2006;79(19):12515–27.

10. Visseaux B, Damond F, Matheron S, Descamps D, Charpentier C. HIV-2 molecular epidemiology. Infect Genet Evol. 2016;46:233–40.

11. Torian L V, Selik RM, Branson B, Owen SM, Granade T, Shouse L, et al. HIV-2 Infection Surveillance --- United States, 1987-2009. Morb Mortal Weekley. 2011;60(29):985–8.

12. Popper SJ, Sarr AD, Travers KU, Guèye-Ndiaye A, Mboup S, Essex ME, et al. Lower Human Immunodeficiency Virus (HIV) Type 2 Viral Load Reflects the Difference in

Pathogenicity of HIV-1 and HIV-2. J Infect Dis. 1999;180(4):1116–21.

13. Ekouévi DK, Avettand-Fènoël V, Tchounga BK, Coffie PA, Sawadogo A, Minta D, et al. Plasma HIV-2 RNA according to CD4 count strata among HIV-2-Infected adults in the IeDEA West Africa Collaboration. PLoS One. 2015;10(6):e0129886.

14. Hawes SE, Sow PS, Stern JE, Critchlow CW, Gottlieb GS, Kiviat NB. Lower levels of HIV-2 than HIV-1 in the female genital tract: correlates and longitudinal assessment of viral shedding. AIDS. 2008;22(18):2517–25.

15. Pádua E, Almeida C, Nunes B, Cortes Martins H, Castela J, Neves C, et al. Assessment of mother-to-child HIV-1 and HIV-2 transmission: An AIDS reference laboratory collaborative study. HIV Med. 2009;10(3):182–90.

16. Mellors JW, Rinaldo Jr CR, Gupta P, White RM, Todd JA, Kingsley LA. Prognosis in HIV-1 Infection Predicted by the Quantity of Virus in Plasma. Science (80-). 1996;272(5265):1167–70.

17. Azevedo-Pereira JM, Santos-Costa Q. HIV interaction with human host: HIV-2 As a model of a less virulent infection. AIDS Rev. 2016;18(1):44–53.

18. Marlink R, Kanki P, Thior I, Travers K, Eisen G, Siby T, et al. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. Science (80-). 1994;265(5178):1587–90.

19. Poropatich K, Sullivan DJ, Sullivan DJ. Human immunodeficiency virus type 1 long-term non-progressors : the viral , genetic and immunological basis for disease non-progression. J Gen Virol. 2011;92(Pt 2):247–68.

20. Kong LI, Lee SW, Kappes JC, Parkin JS, Decker D, Hoxie JA, et al. West African HIV-2-related human retrovirus with attenuated cytopathicity. Science (80-). 1988;240(4858):1525–9.

21. Esbjörnsson J, Månsson F, Kvist A, da Silva ZJ, Andersson S, Fenyö EM, et al. Long-term follow-up of HIV-2-related AIDS and mortality in Guinea-Bissau: a prospective open cohort study. Lancet HIV. 2019;6(1):e25–31.

22. Jespersen S, Hønge BL, Esbjörnsson J, Medina C, da Silva Té D, Correira FG, et al. Differential effects of sex in a West African cohort of HIV-1, HIV-2 and HIV-1/2 dually infected patients: men are worse off. Trop Med Int Heal. 2016;21(2):253–62.

23. Marchant D, Neil SJD, McKnight Á. Human immunodeficiency virus types 1 and 2 have different replication kinetics in human primary macrophage culture. J Gen Virol. 2006;87(2):411–8.

24. Uchtenhagen H, Friemann R, Raszewski G, Spetz AL, Nilsson L, Achour A. Crystal structure of the HIV-2 neutralizing fab fragment 7C8 with high specificity to the V3 region of gp125. PLoS One. 2011;6(4):2–9.

25. Rey M, Krust B, Laurent A, Montagnier L, Hovanessian A. Characterization of human immunodeficiency virus type 2 envelope glycoproteins: dimerization of the glycoprotein precursor during processing. J Virol. 1989;63(2):647–58.

26. Sattentau QJ, Weiss RA. The CD4 antigen: Physiological ligand and HIV receptor. Cell. 1988;52(5):631–3.

27. Hirsch VM, Dapolito G, Goeken R, Campbell BJ. Phylogeny and natural history of the primate lentiviruses, SIV and HIV. Curr Opin Genet Dev. 1995;5(6):798–806.

28. Hill CM, Deng H, Unutmaz D, Kewalramani VN, Bastiani L, Gorny MK, et al. Envelope glycoproteins from human immunodeficiency virus types 1 and 2 and simian immunodeficiency virus can use human CCR5 as a coreceptor for viral entry and make direct CD4-dependent interactions with this chemokine receptor. J Virol. 1997;71(9):6296–304.

29. Clapham PR, Blanc D, Weiss RA. Specific cell surface requirements for the infection of CD4-positive cells by human immunodeficiency virus types 1 and 2 and by simian immunodeficiency virus. Virology. 1991;181(2):703–15.

30. McKnight Á, Clapham PR, Weiss RA. HIV-2 and SIV infection of nonprimate cell lines expressing human CD4: Restrictions to replication at distinct stages. Virology. 1994;201(1):8–18.

31. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science (80-). 1996;272(5263):872–7.

32. Alkhatib G, Combadiere C, Broder CC, Feng Y, Kennedy PE, Murphy PM, et al. CC CKR5: A RANTES, MIP-1α, MIP-1β receptor as a fusion cofactor for macrophage-tropic HIV-1. Science (80-). 1996;272(5270):1955–8.

33. Choe H, Farzan M, Sun Y, Sullivan N, Rollins B, Ponath PD, et al. The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. Cell. 1996;85(7):1135–48.

34. Borrego P, Taveira N. HIV-2 susceptibility to entry inhibitors. AIDS Rev. 2013;15(1):49–61.

35. Raymond S, Delobel P, Mavigner M, Cazabat M, Encinas S, Souyris C, et al. CXCR4-using viruses in plasma and peripheral blood mononuclear cells during primary HIV-1 infection and impact on disease progression. AIDS. 2010;24(15):2305–12.

36. Santos-Costa Q, Lopes MM, Calado M, Azevedo-Pereira JM. HIV-2 interaction with cell coreceptors: amino acids within the V1/V2 region of viral envelope are determinant for CCR8, CCR5 and CXCR4 usage. Retrovirology. 2014;11(1):1–17.

37. Endres MJ, Clapham PR, Marsh M, Ahuja M, Turner JD, McKnight Á, et al. CD4-

independent infection by HIV-2 is mediated by Fusin/CXCR4. Cell. 1996;87(4):745–56.

38. Reeves JD, Hibbitts S, Simmons G, McKnight A, Azevedo-Pereira JM, Moniz-Pereira J, et al. Primary human immunodeficiency virus type 2 (HIV-2) isolates infect CD4-negative cells via CCR5 and CXCR4: comparison with HIV-1 and simian immunodeficiency virus and relevance to cell tropism in vivo. J Virol. 1999;73(9):7795–804.

39. Isaka Y, Sato A, Miki S, Kawauchi S, Sakaida H, Hori T, et al. Small amino acid changes in the V3 loop of human immunodeficiency virus type 2 determines the coreceptor usage for CXCR4 and CCR5. Virology. 1999;264(1):237–43.

40. Gallo SA, Reeves JD, Garg H, Foley BT, Doms RW, Blumenthal R. Kinetic studies of HIV-1 and HIV-2 envelope glycoprotein-mediated fusion. Retrovirology. 2006;3(90):1–8.

41. Takeda E, Kono K, Hulme AE, Hope TJ, Nakayama EE, Shioda T. Fluorescent image analysis of HIV-1 and HIV-2 uncoating kinetics in the presence of old world monkey TRIM5α. PLoS One. 2015;10(3):e0121199.

42. Ren J, Bird LE, Chamberlain PP, Stewart-Jones GB, Stuart DI, Stammers DK. Structure of HIV-2 reverse transcriptase at 2.35-A resolution and the mechanism of resistance to non-nucleoside inhibitors. Proc Natl Acad Sci. 2002;99(22):14410–5.

43. Freund, Frédéric; Boulmé, Florence; Litvak, Simon; Tarrago-Litvak L. Initiation of HIV-2 reverse transcription: a secondary structure model of the RNA-tRNALys3 duplex. Nucleic Acids Res. 2002;29(13):2757–65.

44. Kennedy EM, Gavegnano C, Nguyen L, Slater R, Lucas A, Fromentin E, et al. Ribonucleoside triphosphates as substrate of human immunodeficiency virus type 1 reverse transcriptase in human macrophages. J Biol Chem. 2010;285(50):39380–91.

45. Diamond TL, Roshal M, Jamburuthugoda VK, Reynolds HM, Merriam AR, Lee KY, et al. Macrophage tropism of HIV-1 depends on efficient cellular dNTP utilization by reverse transcriptase. J Biol Chem. 2004;279(49):51545–53.

46. Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, et al. SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. Nat Immunol. 2012;13(3):223–8.

47. Lenzi GM, Domaoal RA, Kim DH, Schinazi RF, Kim B. Mechanistic and kinetic differences between reverse transcriptases of Vpx coding and non-coding lentiviruses. J Biol Chem. 2015;290(50):30078–86.

48. Sakai Y, Doi N, Miyazaki Y, Adachi A, Nomaguchi M. Phylogenetic insights into the functional relationship between primate lentiviral reverse transcriptase and accessory proteins Vpx/Vpr. Front Microbiol. 2016;7:1655.

49. Sevilya Z, Loya S, Adir N, Hizi A. The ribonuclease H activity of the reverse transcriptases of human immunodeficiency viruses type 1 and type 2 is modulated by

residue 294 of the small subunit. Nucleic Acids Res. 2003;31(5):1481–7.

50. Popov S, Rexach M, Zybarth G, Railing N, Lee MA, Ratner L, et al. Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. EMBO J. 1998;17(4):909–17.

51. Tristem M, Marshall C, Karpas A, Petrik J, Hill F. Origin of vpx in lentiviruses. Nature. 1990;347(6291):341–2.

52. Tristem M, Purvis A, Quicke DLJ. Complex evolutionary history of primate lentiviral vpr genes. Virology. 1998;240(2):232–7.

53. Belshan M, Ratner L. Identification of the nuclear localization signal of human immunodeficiency virus type 2 Vpx. Virology. 2003;311(1):7–15.

54. Cheng X, Belshan M, Ratner L. Hsp40 Facilitates Nuclear Import of the Human Immunodeficiency Virus Type 2 Vpx-Mediated Preintegration Complex. J Virol. 2007;82(3):1229–37.

55. van Gent DC, Groeneger AA, Plasterk RH. Mutational analysis of the integrase protein of human immunodeficiency virus type 2. Proc Natl Acad Sci U S A. 1992;89(20):9598–602.

56. Van Gent DC, Elgersma Y, Bolk MWJ, Vink C, Plasterk RHA. DNA binding properties of the integrase proteins of human immunodeficiency viruses types 1 and 2. Nucleic Acids Res. 1991;19(14):3821–7.

57. MacNeil A, Sankale J-L, Meloni ST, Sarr AD, Mboup S, Kanki P. Genomic Sites of Human Immunodeficiency Virus Type 2 (HIV-2) Integration: Similarities to HIV-1 In Vitro and Possible Differences In Vivo. J Virol. 2006;80(15):7316–21.

58. Arya SK, Gallo RC. Human immunodeficiency virus type 2 long terminal repeat: analysis of regulatory elements. Proc Natl Acad Sci. 1988;85(24):9753–7.

59. Fenrick R, Malim MH, Hauber J, Le SY, Maizel J, Cullen BR. Functional analysis of the Tat trans activator of human immunodeficiency virus type 2. J Virol. 1989;63(12):5006–12.

60. de Arellano ER, Soriano V, Alcami J, Holguín Á. New findings on transcription regulation across different HIV-1 subtypes. AIDS Rev. 2006;8(1):9–16.

61. Tong-Starksen SE, Welsh TM, Peterlin BM. Differences in transcriptional enhancers of HIV-1 and HIV-2. Response to T cell activation signals. J Immunol. 1990;145(12):4348–54.

62. Leiden JM, Wang CY, Petryniak B, Markovitz DM, Nabel GJ, Thompson CB. A novel Ets-related transcription factor, Elf-1, binds to human immunodeficiency virus type 2 regulatory elements that are required for inducible trans activation in T cells. J Virol.

1992;66(10):5890–7.

63. Soto-Rifo R, Limousin T, Rubilar PS, Ricci EP, Décimo D, Moncorgé O, et al. Different effects of the TAR structure on HIV-1 and HIV-2 genomic RNA translation. Nucleic Acids Res. 2012;40(6):2653–67.

64. de Breyne S, Soto-Rifo R, López-Lastra M, Ohlmann T. Translation initiation is driven by different mechanisms on the HIV-1 and HIV-2 genomic RNAs. Virus Res. 2013;171(2):366–81.

65. L'Hernault A, Greatorex JS, Crowther RA, Lever AML. Dimerisation of HIV-2 genomic RNA is linked to efficient RNA packaging, normal particle maturation and viral infectivity. Retrovirology. 2007;4:90.

66. Luban J, Goff SP. Mutational analysis of cis-acting packaging signals in human immunodeficiency virus type 1 RNA. J Virol. 1994;68(6):3784–93.

67. McCann EM, Lever AM. Location of cis-acting signals important for RNA encapsidation in the leader sequence of human immunodeficiency virus type 2. J Virol. 1997;71(5):4133–7.

68. Ni N, Nikolaitchik OA, Dilley KA, Chen J, Galli A, Fu W, et al. Mechanisms of Human Immunodeficiency Virus Type 2 RNA Packaging: Efficient trans Packaging and Selection of RNA Copackaging Partners. J Virol. 2011;85(15):7603–12.

69. Luo L, Li Y, Kang CY. Expression of gag Precursor Protein and Secretion of Virus-like gag Particles of HIV-2 from Recombinant Baculovirus-Infected Insect Cells. Virology. 1990;179(2):874–80.

70. Saad JS, Ablan SD, Ghanam RH, Kim A, Andrews K, Nagashima K, et al. Structure of the Myristylated Human Immunodeficiency Virus Type 2 Matrix Protein and the Role of Phosphatidylinositol-(4,5)-Bisphosphate in Membrane Targeting. J Mol Biol. 2008;382(2):434–47.

71. Alford JE, Marongiu M, Watkins GL, Anderson EC. Human Immunodeficiency Virus Type 2 (HIV-2) Gag Is Trafficked in an AP-3 and AP-5 Dependent Manner. PLoS One. 2016;11(7):1–16.

72. Doms RW, Earl PL, Chakrabarti S, Moss B. Human immunodeficiency virus types 1 and 2 and simian immunodeficiency virus env proteins possess a functionally conserved assembly domain. JVirol. 1990;64(0022-538X):3537–40.

73. Voss G, Kirchhoff F, Nick S, Moosmayer D, Gelderblom H, Hunsmann G. Morphogenesis of recombinant HIV-2 gag core particles. Virus Res. 1992;24(2):197–210.

74. Malim MH, Bieniasz PD. HIV Restriction Factors and Mechanisms of Evasion. Cold Spring Harb Perspect Med. 2012;2(5):a006940.

75. Kim B, Nguyen LA, Daddacha W, Hollenbaugh JA. Tight interplay among SAMHD1 protein level, cellular dNTP levels, and HIV-1 proviral DNA synthesis kinetics in human primary monocyte-derived macrophages. J Biol Chem. 2012;287:21570–4.

76. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, et al. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. Nature. 2011;474(7353):658–61.

77. Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Ségéral E, et al. SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. Nature. 2011;474(7253):654–7.

78. Chen C-Y, Shingai M, Welbourn S, Martin MA, Borrego P, Taveira N, et al. Antagonism of BST-2/Tetherin Is a Conserved Function of the Env Glycoprotein of Primary HIV-2 Isolates. J Virol. 2016;90(24):11062–74.

79. Cohen ÉA. From arrest to escape: HIV-1 Vpr cuts a deal. Cell Host Microbe. 2014;15(2):125–7.

80. Chougui G, Margottin-Goguet F. HUSH, a link between intrinsic immunity and HIV latency. Front Microbiol. 2019;10(224).

81. Saleh S, Vranckx L, Gijsbers R, Christ F, Debyser Z. Insight into HIV-2 latency may disclose strategies for a cure for HIV-1 infection. J virus Erad. 2017;3(1):7–14.

82. Deeks SG. Shock and kill. Nature. 2012;487(7408):439–40.

83. Archin NM, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Crooks AM, et al. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. Nature. 2012;487(7408):482–5.

84. Tsai A, Irrinki A, Kaur J, Cihlar T, Kukolj G, Sloan DD, et al. Toll-Like Receptor 7 Agonist GS-9620 Induces HIV Expression and HIV-Specific Immunity in cells from HIV-Infected Individuals on Suppressive Antiretroviral Therapy. J Virol. 2017;91(8):1–20.

85. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell. 2013;155(3):540.

86. Shan L, Siliciano RF. From reactivation of latent HIV-1 to elimination of the latent reservoir: The presence of multiple barriers to viral eradication. BioEssays. 2013;35(6):544–52.

87. Laird GM, Siliciano JD, Siliciano RF, Laird GM, Bullen CK, Rosenbloom DIS, et al. Ex vivo analysis identifies effective HIV-1 latency – reversing drug combinations Find the latest version : Ex vivo analysis identifies effective HIV-1 latency – reversing drug combinations. J Clin Invest. 2015;125(5):1901–12.

88. Imamichi H, Dewar RL, Adelsberger JW, Rehm CA, O'Doherty U, Paxinos EE, et al. Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. Proc Natl Acad Sci. 2016;113(31):8783–8.

89. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, et al. Defective proviruses rapidly accumulate during acute HIV-1 infection. Nat Med. 2016;22(9):1043–9.

90. Sharaf R, Lee GQ, Sun X, Etemad B, Aboukhater LM, Hu Z, et al. HIV-1 proviral landscapes distinguish posttreatment controllers from noncontrollers. J Clin Invest. 2018;128(9):4074–85.

91. Rodes B, Toro C, Jimenez V, Soriano V. Viral Response to Antiretroviral Therapy in a Patient Coinfected with HIV Type 1 and Type 2. Clin Infect Dis. 2005;41(2):e19–21.

92. Murray AJ, Bruner KM, Kumar, MR, Timmons AE, Liu P-T, Clements DH, et al. SIV Proviral Landscape Differs from that of HIV-1 and Shows Gross Hypermutation. J Virus Erad. 2017;3(Supplement 5).

93. Bender AM, Simonetti FR, Kumar MR, Fray EJ, Bruner KM, Timmons AE, et al. The Landscape of Persistent Viral Genomes in ART-Treated SIV, SHIV, and HIV-2 Infections. Cell Host Microbe. 2019;26(1):73-85.e4.

94. Chun TW, Engel D, Berrey MM, Shea T, Corey L, Fauci AS. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. Proc Natl Acad Sci U S A. 1998;95(15):8869–73.

95. Chavez L, Calvanese V, Verdin E. HIV Latency Is Established Directly and Early in Both Resting and Activated Primary CD4 T Cells. PLoS Pathog. 2015;11(6):1–21.

96. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, et al. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nat Med. 1999;5(5):512–7.

97. Chun T-W, Carruth L, Finzi D, Shen X, Digiuseppe JA, Taylor H, et al. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature. 1997;387(6629):183–8.

98. Lee GQ, Yu XG, Lichterfeld M, Lee GQ, Orlova-fink N, Einkauf K, et al. Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4 + T cells Find the latest version : Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4 + T cells. J Clin Invest. 2017;127(7):2689–96.

99. Hiener B, Horsburgh BA, Eden JS, Barton K, Schlub TE, Lee E, et al. Identification of Genetically Intact HIV-1 Proviruses in Specific CD4+ T Cells from Effectively Treated Participants. Cell Rep. 2017;21(3):813–22.

100. Vanhamel J, Bruggemans A, Debyser Z. Establishment of latent HIV-1 reservoirs: what

do we really know? J virus Erad. 2019;5(1):3–9.

101. Whitney JB, Hill AL, Sanisetty S, Penaloza-Macmaster P, Liu J, Shetty M, et al. Rapid Seeding of the Viral Reservoir Prior to SIV Viremia in Rhesus Monkeys. Nature. 2014;512(7512):74–7.

102. Steingrover R, Pogány K, Fernandez Garcia E, Jurriaans S, Brinkman K, Schuitemaker H, et al. HIV-1 viral rebound dynamics after a single treatment interruption depends on time of initiation of highly active antiretroviral therapy. Aids. 2008;22(13):1583–8.

103. Persaud D, Gay H, Ziemniak C, Chen YH, Piatak M, Chun T-W, et al. Absence of Detectible HIV-1 Viremia After Treatment Cessation in an Infant. N Engl J Med. 2013;369(19):1828–35.

104. Ananworanich J, Robb ML. The transient HIV remission in the Mississippi baby: why is this good news? J Int AIDS Soc. 2014;17:19859.

105. Li B, Gladden AD, Altfeld M, Kaldor JM, Cooper DA, Kelleher AD, et al. Rapid Reversion of Sequence Polymorphisms Dominates Early Human Immunodeficiency Virus Type 1 Evolution. J Virol. 2006;81(1):193–201.

106. Reeves JD, Doms RW. Human immunodeficiency virus type 2. J Gen Virol. 2002;83:1253–65.

107. Hiener B, Eden J-S, Horsburgh BA, Palmer S. Amplification of Near Full-length HIV-1 Proviruses for Next-Generation Sequencing. J Vis Exp. 2018;(140):1–12.

108. Los Alamos National Laboratory. HIV Sequence Database – Hypermut tool.

109. Los Alamos National Laboratory. HIV Sequence Database – HIVAlign tool.

110. Geneious Prime 2019.1.1.

111. Campbell-Yesufu OT, Gandhi RT. Update on human immunodeficiency virus (HIV)-2 infection. Clin Infect Dis. 2011;52(6):780–7.

112. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. Nat Med. 2009;15(8):893–900.

113. Pollack RA, Jones RB, Pertea M, Bruner KM, Martin AR, Thomas AS, et al. Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. Cell Host Microbe. 2017;21(4):494-506.e4.

114. Kuzembayeva M, Dilley K, Sardo L, Hu WS. Life of psi: How full-length HIV-1 RNAs become packaged genomes in the viral particles. Virology. 2014;454–455:362–70.

115. Gottlieb GS, Raugi DN, Smith RA. 90-90-90 for HIV-2? Ending the HIV-2 epidemic by

enhancing care and clinical management of patients infected with HIV-2. Lancet HIV. 2018;5(7):e390-99.

116. de Silva TI, Cotten M, Rowland-Jones SL. HIV-2: the forgotten AIDS virus. Trends Microbiol. 2008;16(12):588–95.

117. MacNeil A, Sarr AD, Sankale J-L, Meloni ST, Mboup S, Kanki P. Direct Evidence of Lower Viral Replication Rates In Vivo in Human Immunodeficiency Virus Type 2 (HIV-2) Infection than in HIV-1 Infection. J Virol. 2007;81(10):5325–30.

118. Hamer D. Can HIV be Cured? Mechanisms of HIV Persistence and Strategies to Combat It. Curr HIV Res. 2004;2(2):99–111.

119. Savarino A, Mai A, Norelli S, El Daker S, Valente S, Rotili D, et al. "Shock and kill" effects of class I-selective histone deacetylase inhibitors in combination with the glutathione synthesis inhibitor buthionine sulfoximine in cell line models for HIV-1 quiescence. Retrovirology. 2009;6(52):1–10.

120. Archin NM, Margolis DM. Emerging strategies to deplete the HIV reservoir. Curr Opin Infect Dis. 2014;27(1):29–35.

121. Deeks SG, Lewin SR, Ross AL, Ananworanich J, Benkirane M, Cannon P, et al. International AIDS Society global scientific strategy: Towards an HIV cure 2016. Nat Med. 2016;22(8):839–50.

122. Abner E, Jordan A. HIV "shock and kill" therapy: In need of revision. Antiviral Res. 2019;166(November 2018):19–34.

123. Kim Y, Anderson JL, Lewin SR. Getting the "Kill" into "Shock and Kill": Strategies to Eliminate Latent HIV. Cell Host Microbe. 2018;23(1):14–26.

124. Cillo AR, Sobolewski MD, Bosch RJ, Fyne E, Piatak M, Coffin JM, et al. Quantification of HIV-1 latency reversal in resting CD4+ T cells from patients on suppressive antiretroviral therapy. Proc Natl Acad Sci U S A. 2014;111(19):7078–83.

125. Archin NM, Bateson R, Tripathy MK, Crooks AM, Yang KH, Dahl NP, et al. HIV-1 expression within resting CD4+ T cells after multiple doses of vorinostat. J Infect Dis. 2014;210(5):728–35.

126. Elliott JH, Wightman F, Solomon A, Ghneim K, Ahlers J, Cameron MJ, et al. Activation of HIV Transcription with Short-Course Vorinostat in HIV-Infected Patients on Suppressive Antiretroviral Therapy. PLoS Pathog. 2014;10(10):e1004473.

127. Wei DG, Chiang V, Fyne E, Balakrishnan M, Barnes T, Graupe M, et al. Histone Deacetylase Inhibitor Romidepsin Induces HIV Expression in CD4 T Cells from Patients on Suppressive Antiretroviral Therapy at Concentrations Achieved by Clinical Dosing. PLoS Pathog. 2014;10(4).

128. Rasmussen T, Tolstrup M, Brinkmann C, Olesen R, Erikstrup C, Solomon A, et al. Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. Lancet HIV. 2014;1(1):e13-21.

129. Søgaard OS, Graversen ME, Leth S, Olesen R, Brinkmann CR, Nissen SK, et al. The Depsipeptide Romidepsin Reverses HIV-1 Latency In Vivo. PLoS Pathog. 2015;11(9):e1005142.

130. Bullen CK, Laird GM, Durand CM, Siliciano JD, Siliciano RF. New ex vivo approaches distinguish effective and ineffective single agents for reversing HIV-1 latency in vivo. Nat Med. 2014;20(4):425–9.

131. Cary DC, Fujinaga K, Peterlin BM. Molecular mechanisms of HIV latency. J Clin Invest. 2016;126(2):448–54.

132. Mohammadi P, di Iulio J, Muñoz M, Martinez R, Bartha I, Cavassini M, et al. Dynamics of HIV Latency and Reactivation in a Primary CD4+ T Cell Model. PLoS Pathog. 2014;10(5):e1004156.

133. Jones RB, Mueller S, O'Connor R, Rimpel K, Sloan DD, Karel D, et al. A Subset of Latency-Reversing Agents Expose HIV-Infected Resting CD4 + T-Cells to Recognition by Cytotoxic T-Lymphocytes. PLoS Pathog. 2016;12(4):e1005545.

134. Gutiérrez C, Serrano-Villar S, Madrid-Elena N, Pérez-Eĺas MJ, Martń ME, Barbas C, et al. Bryostatin-1 for latent virus reactivation in HIV-infected patients on antiretroviral therapy. Aids. 2016;30(9):1385–92.

135. Novis CL, Archin NM, Buzon MJ, Verdin E, Round JL, Lichterfeld M, et al. Reactivation of latent HIV-1 in central memory CD4+ T cells through TLR-1/2 stimulation. Retrovirology. 2013;10(119).

136. Schlaepfer E, Speck RF. TLR8 Activates HIV from Latently Infected Cells of Myeloid-Monocytic Origin Directly via the MAPK Pathway and from Latently Infected CD4 + T Cells Indirectly via TNF-α. J Immunol. 2011;186(7):4314–24.

137. Offersen R, Nissen SK, Rasmussen TA, Østergaard L, Denton PW, Søgaard OS, et al. A Novel Toll-Like Receptor 9 Agonist, MGN1703, Enhances HIV-1 Transcription and NK Cell-Mediated Inhibition of HIV-1-Infected Autologous CD4 + T Cells. J Virol. 2016;90(9):4441–53.

138. Lim S, Osuna CE, Hraber PT, Hesselgesser J, Gerold JM, Barnes TL, et al. TLR7 agonists induce transient viremia and reduce the viral reservoir in SIV-infected rhesus macaques on antiretroviral therapy. 2018;10(439):eaao4521.

139. Policicchio BB, Xu C, Brocca-Cofano E, Raehtz KD, He T, Ma D, et al. Multi-dose Romidepsin Reactivates Replication Competent SIV in Post- antiretroviral Rhesus Macaque Controllers. PLoS Pathog. 2016;12(9):e1005879.

140. Del Prete GQ, Oswald K, Lara A, Shoemaker R, Smedley J, Macallister R, et al. Elevated Plasma Viral Loads in Romidepsin-Treated Simian Immunodeficiency Virus-Infected Rhesus Macaques on Suppressive. Antimicrob Agents Chemother. 2016;60(3):1560–72.

141. Gama L, Abreu CM, Shirk EN, Price SL, Li M, Laird GM, et al. Reactivation of simian immunodeficiency virus reservoirs in the brain of virally suppressed macaques. Aids. 2017;31(1):5–14.

142. Xing S, Siliciano R. Targeting HIV latency: pharmacologic strategies toward eradication. Drug Discov Today. 2013;18(0):541–51.

143. Shirakawa K, Chavez L, Hakre S, Calvanese V, Verdin E. Reactivation of latent HIV by histone deacetylase inhibitors. Trends Microbiol. 2013;21(6):277–85.

144. Walker-Sperling VEK, Cohen VJ, Tarwater PM, Blankson JN. Reactivation Kinetics of HIV-1 and Susceptibility of Reactivated Latently Infected CD4 + T Cells to HIV-1-Specific CD8 + T Cells . J Virol. 2015;89(18):9631–8.

145. Bosque A, Planelles V. Induction of HIV-1 latency and reactivation in primary memory CD4 + T cells Induction of HIV-1 latency and reactivation in primary memory CD4 ♀ T cells. Blood. 2011;113(1):58–65.

146. Avettand-Fenoel V, Damond F, Gueudin M, Matheron S, Mélard A, Collin G, et al. New sensitive one-step real-time duplex PCR method for group A and B HIV-2 RNA load. J Clin Microbiol. 2014;52(8):3017–22.

147. Malnati MS, Scarlatti G, Gatto F, Salvatori F, Cassina G, Rutigliano T, et al. A universal real-time PCR assay for the quantification of group-M HIV-1 proviral load. Nat Protoc. 2008;3(7):1240–8.

148. Lopatin U, Wolfgang G, Tumas D, Frey CR, Ohmstede C, Hesselgesser J, et al. Safety, pharmacokinetics and pharmacodynamics of GS-9620, an oral Toll-like receptor 7 agonist. Antivir Ther. 2013;18:409–18.

149. Smith RL, Pizer LI, Johnson EM, Wilcox CL. Activation of second-messenger pathways reactivates latent herpes simplex virus in neuronal cultures. Virology. 1992;188:311–8.

150. Siekevitz M, Josephs SF, Dukovich M, Peffer N, Wong-staal F, Greene WC. Activation of the HIV-1 LTR by T Cell Mitogens and the Trans-Activator Protein of HTLV- 1. Science (80-). 1987;238(4833):1575–8.

151. Chatila T, Silverman L, Miller R, Geha R. Mechanisms of T cell activation by the calcium ionophore ionomycin. J Immunol. 1989;143(4):1283–9.

152. Kiselinova M, Spiegelaere WD, Buzon MJ, Malatinokova E, Lichterfeld M, Vandekerckhove L. Integrated and Total HIV-1 DNA Predict Ex Vivo Viral Outgrowth. PLoS Pathog. 2016;12(3):1–17.

153. Guedin M, Damond F, Braum J, Taïeb A, Lemée J, Plantier JC, et al. Differences in proviral DNA load between HIV-1- and HIV-2-infected participants. AIDS. 2008;22(2):211–5.

154. Gottlieb GS, Hawes SE, Kiviat NB, Sow PS. Differences in proviral DNA load between HIV-1-infected and HIV-2-infected patients. AIDS. 2008;22(11):1379–80.

155. Abram ME, Ferris AL, Shao W, Alvord WG, Hughes SH. Nature, Position, and Frequency of Mutations Made in a Single Cycle of HIV-1 Replication. J Virol. 2010;84(19):9864–78.

156. Lecossier D, Bouchonnet F, Clavel F, Hance AJ. Hypermutation of HIV-1 DNA in the absence of the Vif protein. Science (80-). 2003;300(5622):1112.

157. Laboratory LAN. HIV Sequence Database - Gene Cutter tool.

158. Frenkel LM, Wang Y, Learn GH, McKernan JL, Ellis GM, Mohan KM, et al. Multiple Viral Genetic Analyses Detect Low-Level Human Immunodeficiency Virus Type 1 Replication during Effective Highly Active Antiretroviral Therapy. J Virol. 2003;77(10):5721–30.

159. Kearney MF, Wiegand A, Shao W, McManus WR, Bale MJ, Luke B, et al. Ongoing HIV Replication During ART Reconsidered. Open Forum Infect Dis. 2017;4(3):1–5.

160. Taube R, Peterlin BM. Lost in transcription: Molecular mechanisms that control HIV latency. Viruses. 2013;5(3):902–27.

161. Lusic M, Marcello A, Cereseto A, Giacca M. Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter. EMBO J. 2003;22(24):6550–61.