MOLECULAR MECHANISMS AND HOST FACTORS INVOLVED IN HIV-1 LATENCY

Mrudhula Madapuji Srinivasan

Thesis submitted to the

University of Ottawa

in partial fulfillment of the requirements for the

Master of Science degree in Microbiology and Immunology

Department of Biochemistry, Microbiology, and Immunology

Faculty of Medicine

University of Ottawa

ABSTRACT

The Human Immunodeficiency virus-1 can stay undetected and unaffected by host immune surveillance and antiretroviral therapy. This phenomenon is called proviral latency and the cells harbouring such viruses are part of the latently infected cell reservoir. In this situation, the viral genome integrates into the host's genome upon infection, whereby infected cells exhibit either very low levels or no viral transcription, and hence no viral proteins or egress viruses are produced that can be detected by the immune system. However, viral transcription can be re-activated to produce infectious viruses under certain circumstances. Host-encoded retroviral restriction factors like APOBEC3 (A3) proteins are part of our intrinsic immune defences against retroviral infection, introducing mutations in viral replication intermediates. We hypothesize that low levels of G-to-A transition mutations in the HIV-1 LTR region, introduced by APOBEC3G/F, could lead to a latency-like phenotype. These latent viruses pose major hurdles for HIV-1 cure therapies. Our lab previously created a library of clones possessing mutations in the LTR introduced by A3G/F. Later, mutated LTRs were cloned into 3 types of plasmid backbones: 1) a pEGFP expression vector to study the transcriptional activity of the mutated promoter, 2) into non-replicative pNL4 Δ*env* Δ*vif* viral expression vector, and 3) into a replicative pNL4-CXCR4 viral vector to study infection and induction by latency reversal agent (LRA) treatment to better understand the mechanism of latency and transcriptional induction. Viruses produced from these plasmids carrying mutated promoters are referred to as latency-prone viruses or LPVs in this thesis. Characterizing the transcription, infection, and induction to PMA/I of the LPVs would essentially help in evaluating the role of A3 mutations in viral latency and further help in the development of new therapeutics.

ACKNOWLEDGEMENTS

I would like to thank Dr. Marc-André Langlois for giving me the opportunity to work in his laboratory as a master's student. I would also like to thank Dr. Martin Pelchat for his role as my co-supervisor. I am thankful to Dr. Samar Dankar for her role in this project. I also want to thank Matthew Greig and Cindy Lam for generating the LPV library and conducting the early experimental troubleshooting. I would also like extend my thanks to the current members of the Langlois lab for their inputs during lab meetings.

LIST OF ABBREVIATIONS

AIDS: Acquired Immunodeficiency Syndrome

AP-1: Activator protein 1

APOBEC3/A3: Apolipoprotein B mRNA editing enzyme, catalytic protein-like 3

ART: Anti-retroviral therapy

BRD4: bromodomain protein 4

CD3: Cluster of differentiation 3

CD4: Cluster of differentiation 4

CDK9: Cyclin dependent kinase K9

CycT1: Cyclin T1

CCR5: C-C chemokine receptor type 5

CHIP: Chromatin Immunoprecipitation

CXCR4: C-X-C chemokine receptor type 4

DRB: 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole

DSIF: DRB sensitivity-inducing factor

E. coli: *Escherichia coli*

Env: Envelope protein

Gag: Group-specific antigen

GALT: Gut-associated lymphoid tissues

HDAC: Histone deacetylase

HEXIM1: hexamethylene bisacetamide inducible protein 1

HIV-1: Human Immunodeficiency Virus Type 1

HMT: Histone methyltransferase

HPCs: Haematopoietic stem cells

IN: Integrase enzyme

LARP7: La-related protein

LPV: Latency-prone virus

LR: Latent reservoir

LRA: Latency reversing agent.

LTR: Long terminal repeat

MePCE: 5'methylphosphate capping enzyme

MFI: Mean fluorescence intensity

mRNA: Messenger ribonucleic acid

Nef: Negative regulatory factor

NELF: negative elongation factor

NFAT: Nuclear factor of activated T-cells

NF-κB: Nuclear factor kappa-light-chain enhancer of activated B cells

ORF: Open reading frame

PBMC: Peripheral blood mononuclear cells

PKC: Protein kinase C

PLWH: People living with HIV.

PMA/I: Phorbol 12 myristate 13 acetate/ Ionomycin

Pol: Polymerase

P-TEFb: positive transcription elongation factor b

Rev: Regulator of expression of virion proteins

RNA Pol II: RNA Polymerase II

RT: Reverse Transcriptase

snRNP: small nuclear Ribonucleoprotein

Sp-1: Specificity protein 1

STAT5: Signal transducer and activator of transcription 5

TAR: Trans-activation response element

Tat: Trans-activator of transcription

TFBS: Transcription factor binding site.

Vif: Viral infectivity factor

VOAs: Viral outgrowth assays

Vor: Vorinostat

Vpr: Viral protein R

Vpu: Viral protein U

VSV-G: Vesicular Stomatitis virus glycoprotein

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

HIV-1, which stands for Human Immunodeficiency virus type 1 (addressed as HIV thereafter), was first discovered in the year 1983 and was later found to be associated with the 1984 AIDS (acquired immunodeficiency syndrome) pandemic [1, 2]. HIV belongs to the subfamily Orthoretrovirinae of the family Retroviridae [1, 2]. The family Retroviridae includes enveloped RNA viruses that make use of an enzyme called reverse transcriptase to convert their RNA into a DNA intermediate [3]. This intermediate would then integrate into the host's genome with the help of the integrase enzyme to function as a part of the host's cellular DNA. The virus replicates thereafter using the host cell's transcriptional machinery [1-3].

1.1 HIV structure and genome

The HIV virus has a conical capsid core housed within the plasma membrane that are embedded with viral envelope glycoproteins [1-3]. The viral capsid core houses two copies of a single-stranded RNA genome (**Figure 1.1**). Each viral RNA genome is 9.7 kb long with a diameter of 120 nm [4]. The viral RNA is protected from nucleases by the tightly bound nucleocapsid proteins p6 and p7. Reverse transcriptase, integrase, and protease are the major enzymes for viral survival post-infection and are a present as a part of the viral core along with the viral RNA [4]. Capsid proteins (p24) form the icosahedral capsid. Matrix proteins (p17) make the underlayer of the viral envelope. Viral proteins Vif, Vpr, and Nef are also present within the virion particle [5]. When the capsid egresses from the host cell, it takes a portion of the host-cell membrane with it thereby creating the envelope. The transmembrane glycoprotein (gp41) and the exterior surface glycoprotein (gp120), which together make up the HIV spikes, are embedded within the lipid bilayer [6].

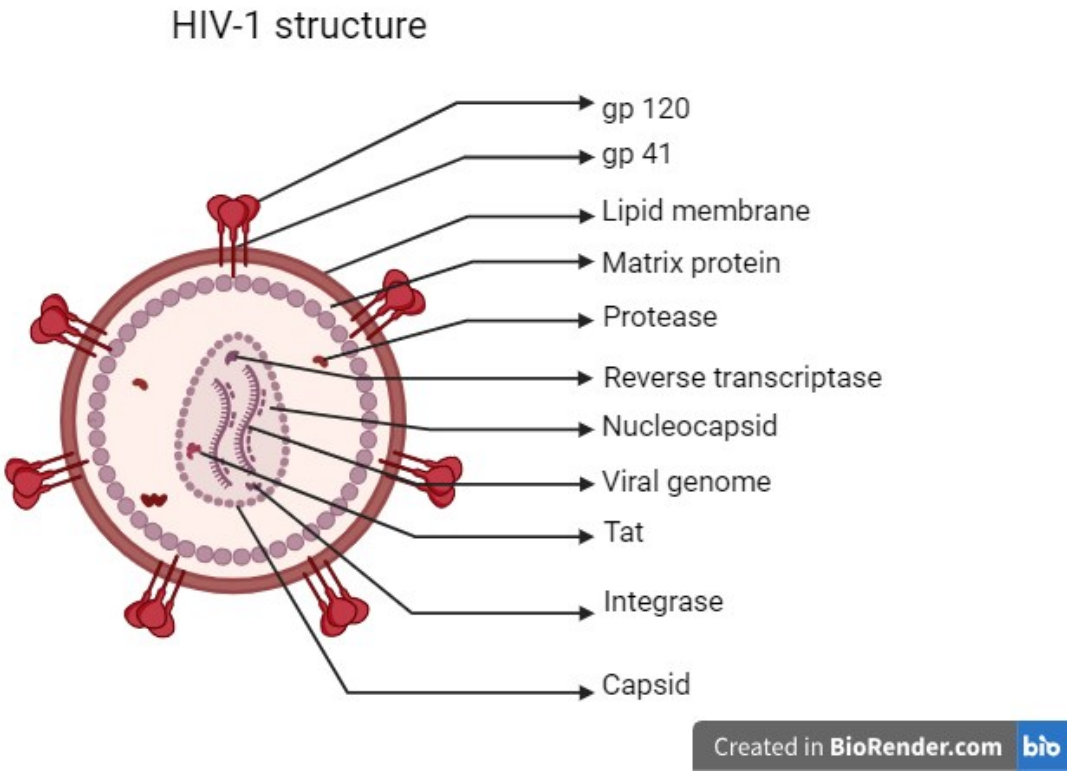


Figure 1.1 Structure of HIV virus
Created using Biorender.

The HIV viral genome has 9 open reading frames (ORF) encoding for 15 proteins (**figure 1.2 A**) [7-9]. The three genes *gag*, *pol*, and *env* are expressed as three different polyproteins. The polyprotein Gag precursor is cleaved by protease to generate matrix (MA), capsid (CA) and Nucleocapsid (NC) (**Figure 1.2 A**) [7, 9]. The Pol (polymerase) polyprotein cleavage results in protease (PR), reverse transcriptase (RT), and integrase (IN) [7, 9]. The Env (envelope) polyprotein is cleaved into glycoproteins gp120 and gp41 [7, 9, 10]. Furthermore, there are six accessory proteins which make-up the total of 15 proteins; transactivator of transcription (Tat), negative regulatory factor (Nef), viral infectivity factor (Vif), viral protein u (Vpu), viral protein r (Vpr) and regulator of expression of viral proteins (Rev) [9]. The HIV genome is flanked by long terminal repeats (LTR) on both ends. 5'LTR serves as the promotor for viral transcription. The 3'LTR is also crucial because it is copied to the 5'end of the virus during reverse transcription [11].

The LTRs have 3 main regions: U3, R, and U5 (**refer figure 1.2 B**) (in 5' – 3'direction). U3 region contains the core promoter, TATA box, enhancer region and modulatory regions so, it is very important for viral transcription initiation [3]. This region houses binding sites for transcription factors AP-1, NFAT, NF-κB, Sp-1 and STAT-5 [12-14]; that are important for viral transcription. The R region has the transactivation response element (TAR) which forms an RNA stem loop facilitating its binding to the viral protein Tat. Tat binding TAR region increases the viral transcription level there by switching from promoter proximal pause and helping with transcription elongation [15]. The U5 region has the primer binding site, and primer initiates reverse transcription. It has fewer transcription factor binding sites than the rest of the LTR with only NFAT and Ap-1 binding sites [3]. NFAT, NF-κB and Ap-1 transcription factors are induced or

activated following treatments with protein kinase-C agonists [16]. This is the basis for treatments with latency reversal agents in this thesis.

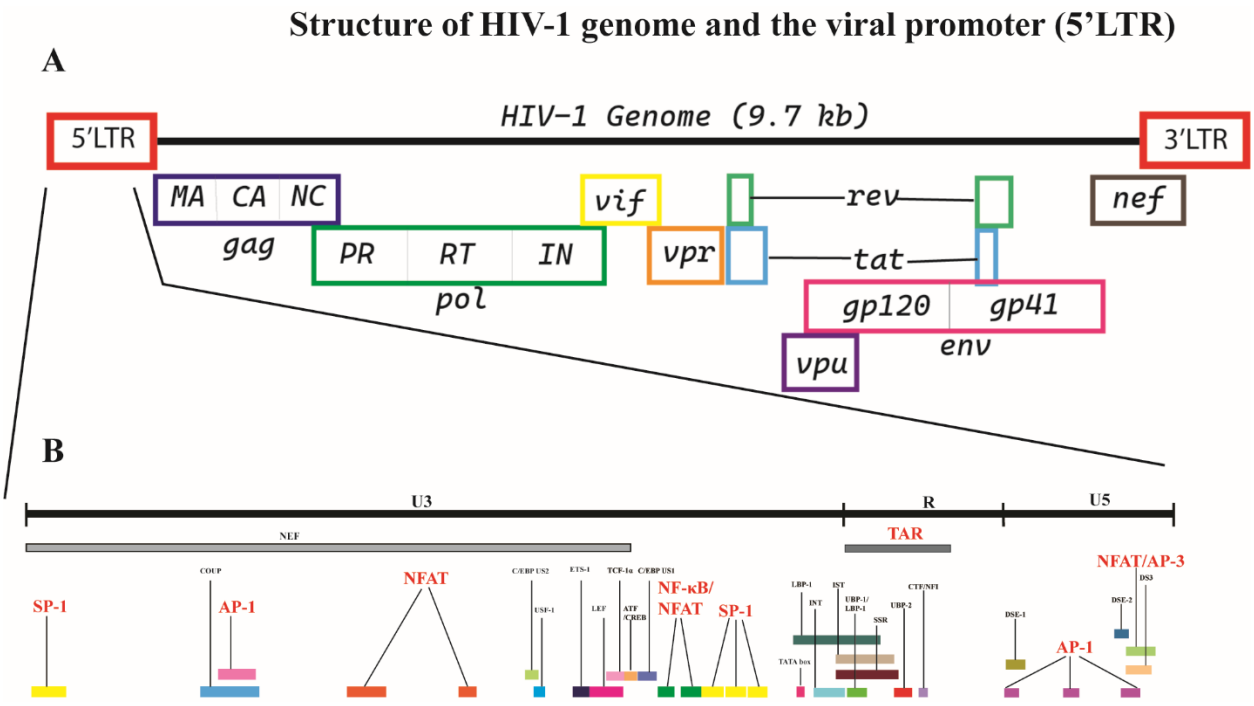


Figure 1.2 Structure of HIV genome and viral promoter

A- HIV genome and its open reading frames.

B- HIV LTR with its important transcription factor binding sites and TAR region.

1.2 HIV latency

HIV latency is an important contributor to the persistence of HIV infection [17]. The persisting HIV infected cells are usually found in latent reservoirs that are mostly composed of long-lived population of resting CD4+ T cells harbouring transcriptionally silent and intact proviruses that retain the ability to produce infectious viruses [18]. The two main types of HIV reservoirs are anatomical and cellular characterized by tissue / organ harbouring HIV infected cells [19, 20].

1.2.1 Clinical latency

An individual infected with HIV undergoes an asymptomatic stage following an acute infection which is described as clinical latency [19]. During this stage, the patient displays no disease-associated symptoms although the virus is undergoing basal replication. This stage is followed by the final or AIDS stage that is characterized by a decline in CD4+ T cell counts and high viral titres in blood. ART (Antiretroviral therapy) administration can prolong clinical latency by a few decades in some patients. In some people living with HIV (PLWH), the viral load is suppressed to undetectable levels with high CD4+ T cell counts even in the absence of ART. These are a rare group of people called elite controllers, however they are not completely rid of the virus and can have active replication if the virus evolves [21]. As of now, (October 2023) there have been three cases confirmed to have been cured of HIV, the Berlin Patient, the London Patient and Düsseldorf patient, had progressive cancerous malignancies that provided the opportunity to perform a bone marrow transplant to treat both their cancer and HIV [22-25]. These patients were transplanted with bone marrow cells from donors having a deletion in the C-C chemokine receptor type 5 (CCR5) (cell-surface co-receptor for HIV) expressing gene, meaning their ‘new’ immune system would be protected from R5-tropic HIV infection [22-25]. In the year 2020, a Brazilian man called as São Paulo patient was regarded as the first person to have been putatively cured of HIV infection

using an easily accessible and well-tolerated drug cocktail [26]. Three more cases of putative HIV cure have been addressed in the article cited, among which one patient was transplanted with cells expressing wildtype CCR5 and still attained HIV cure [25]. In most of these cases, the size of the latent reservoir was decreasing even without ART.

In the overall HIV latent reservoir, only a small subset of the integrated proviruses are capable of producing productive infection because of possible mutational defects in the genome of egress viruses [27]. There are a wide range of defects possibly associated with latency. All infected cells and tissues contain various forms of persisting viruses that may or may not participate in HIV pathogenesis and latency [28]. This also includes defective proviruses which indirectly participate in HIV pathogenesis by their basal viral transcription and synthesis of initial viral proteins [27]. These proteins can induce and maintain immune activation, thus participating in chronic inflammation and also in persistence of HIV pathogenesis [28].

1.2.2 Latent Reservoir

There are two main types of latent reservoirs associated with HIV infection including anatomical reservoirs and cellular reservoirs as discussed earlier. **Figure 1.3** shows the various cellular and anatomical HIV reservoirs commonly studied in literatures.

HIV latency has been mostly associated with CD4+ T cells, partially due to the fact that memory CD4+ T cells have a long lifespan of up to several years. However, HIV infection is not limited to infecting CD4+ T cells, as it was also discussed that monocyte-derived CD4+ cells also play important roles in HIV persistence. The latent monocytes are particularly studied in infection of the central nervous system due to their ability to cross the blood–brain barrier [29, 30]. Moreover, hematopoietic stem cells (HPCs) were also reported to have a controversial role in HIV infection and transmission, as some studies show that HIV is capable of infecting HPCs [14–16], whereas

some others show that HPCs are resistant to HIV infection [17–19]. They have both used *in-vivo* and *ex-vivo* models to demonstrate the ability or inability of HIV to infect HPCs. This indicates that further investigation is needed to clarify infection of HPCs by HIV.

The first type of latent reservoirs are the anatomical reservoirs that serve as hotspots for the spread of HIV within the host lymphoid tissue, mainly in the gut and lymph nodes [31]. Highest viral replication was observed in lymphoid tissues such as lymph nodes, spleen, gut-associated lymphoid tissues (GALT) and thymus during peak infection [32]. Studies conducted with premortem, and postmortem HIV-infected patient tissue samples showed that genital tract tissues including prostate, seminal vesicles and testis were potential sources of virus. Evidence of intact full-length (FL) env proviruses in the frontal and occipital lobes of the brain was also observed [31]. HIV-infected cells were also found in the CNS, lungs, kidneys, liver, adipose tissue, genitourinary tract and bone marrow in various studies [31-33].

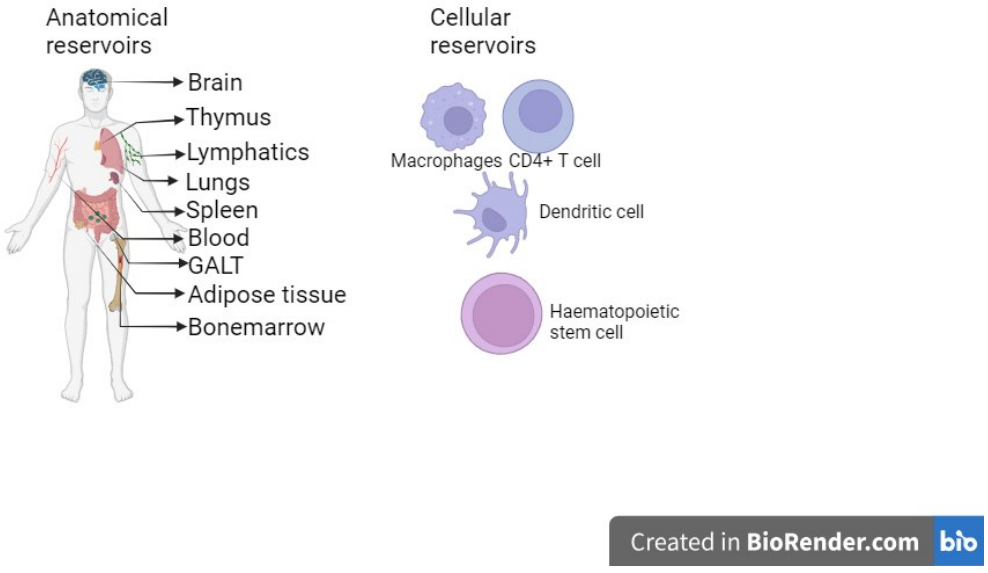


Figure 1.3 HIV latent reservoir
Figure shows various reservoirs of HIV. The right panel describes the various anatomical reservoirs, and the left panel shows the various cellular reservoirs. This figure was created using Biorender.

1.2.3 Half-life of the latent reservoir

ART constitutes the use of potent inhibitors of HIV replication and infection to help in achieving a rapid exponential decline in virus levels in plasma. This is reflected in short half-lives of free viruses (less than 6 hours) and the viruses produced from infected cells (about a day). A few weeks post-ART, sees production of viruses that have a much slower viral decay with half-life of about 14 days. This mainly signifies the chronically infected resting memory CD4+ T cells, persisting in the reservoirs with a much slower turnover [34], despite antiretroviral therapy [35, 36]. The mean half-life of these latent reservoirs is 43.9 months and even if it is assumed that the reservoir has only 1×10^5 cells, their eradication could take up to 60 years. Resting CD4+ T cells when infected latently with HIV, allows HIV to persist in them over a long period of time even when the patients are on effective ART [34]. These cells do not produce virus particles while in resting state but can give rise to infectious virus particles following activation by several stimuli, leading to a viral rebound when ART is interrupted [34, 35, 37]. For example, in normal CD4+ T cell homeostasis, memory T cell clones that proliferate in response to IL-7 cytokine also aid in maintenance of the HIV reservoir [38]. The proliferation of these cells occurs in the absence of any viral reactivations [39, 40], indicating the requirement of low-level proliferation in a memory T cell subset, maintaining the reservoir while hiding from the immune system and evading antiviral treatments [41].

HIV latency can be epigenetic or non-epigenetic (**Figures 1.4 and 1.5**). The various epigenetic and non-epigenetic mechanisms of HIV latency are elaborated in the section below. In addition, we are also discussing what is viral molecular latency, a new concept that we are proposing.

1.3 Epigenetic Mechanisms

There are multiple ways that latency can be established before and after proviral integration such as LTR circle, DNA methylation, repressive histone marks, promoter occlusion and convergent transcription. (**refer figure 1.4**)

1.3.1 Pre-integration latency- LTR circle

The unintegrated viral DNA could be found in linear, 1-LTR and 2-LTR circular forms. Unintegrated linear viral DNA has a very short half-life of up to several days [42], hence it is not potentially contributing to the viral latency. Despite its fast turnover, these linear viral DNA may form truncated or internally rearranged 1-LTR and 2-LTR circular DNA by auto-integration [43, 44] or simple ligation [45]. Whereas, circular unintegrated DNAs are more stable but their role in latency is controversial. Few studies have reported that the circular viral DNAs cannot sustain replication and therefore they have been considered as dead-end products of abortive infections [46], on the other hand, it was also previously shown that unintegrated HIV DNA was organized into some type of chromatin structures coupled with histone modifications that is usually associated with transcriptionally silenced chromatin [47]. A number of studies have also demonstrated that unintegrated HIV DNA have limited expression of Nef [48, 49], but, it was interesting to note that a few articles that studied unintegrated HIV DNA showed that HIV Vpr enhanced Nef expression in these genome [50, 51]. While the role of unintegrated HIV DNA in viral infection is still a matter of controversy, it was shown that the 2-LTR circles of HIV DNA can auto integrate into chromosomal DNA, leading to active infections [52, 53].

Despite the controversial role of unintegrated HIV in latency, the integrated proviruses are still believed to be the only possible source of HIV infection and protein expression, as post-integration HIV latency is usually established very early on during acute infection and produces stable and long lived proviral reservoirs [53]. The following sections will discuss a variety of factors that induce post-integration latency.

1.3.2 The post-integration latency results from

1. The repressive nucleosome array on the HIV promoter located in the 5′LTR

The positioning of nucleosomes has been shown to hinder HIV replication leading it to latency [54]. The precise organisation of nucleosomes nuc-0 and nuc-1 in HIV 5' LTR with respect to the cis-acting regulatory elements is vital for viral replication [54]. Nuc-1 is usually located near the transcription start site and its displacement during transcriptional activation highlights its central role in regulation of HIV transcription [55]. The disruption of these nucleosomes from their respective positions will lead to loss of viral transcription and hence viral latency.

2. The accumulation of histone repressive marks

Histones modifications are capable of altering the state of HIV latency [29]. Even though DNA methylation has suppressive functions, histone methylation may result in activation or suppression of gene expression. Their impact on gene expression depends on the site of methylation [56]. Histone acetylation, on the other hand, is normally associated with active transcription [57].

Lysine and arginine residues that are abundant in histones, are specifically methylated by histone methyltransferases [29]. The promoters of suppressed genes were frequently found to have trimethylation of histone H3 at their lysine 9 (H3K9me3). Two histone methyltransferases, SUV39H1 and G9a, are known to mediate this modification. Interestingly, both methyltransferases maintain HIV latency, as experimental knockdown of both of these proteins was demonstrated to

reactivate HIV in latently infected cell lines [58]. H3K27me3 is also an epigenetic marker of silenced genes observed on the silenced HIV LTRs, its demethylation sensitises the latent proviruses to external stimuli that can reactivate the virus [59]. Knockdown of the methyltransferase EZH2 also induces demethylation of H3K27 and therefore induces a similar effect, sensitizes latent proviruses to external stimuli that reactive latent proviruses [59].

3. DNA methylation of two CpG islands surrounding the transcription start site (TSS)

In resting CD4+ T cells, multiple factors aid in maintenance of HIV latency by inhibiting viral gene expression after integration with the host cellular DNA. We know that, chromatin structure present at the integration site of the provirus affects the proviral transcription [60] and the several deep analyses done on the integrated HIV proviruses have concluded that latent proviruses are often presented with unique epigenetic patterns that are not found in normal proviruses. For instance, methylation of cytosine on CpG-rich promoters is a marker of transcriptional silencing commonly found in mammalian genome [61]. The same applies to HIV provirus, as studies have shown that their CpG regions to be methylated. DNA methylation restricts the transcription factors from accessing the HIV promoters [62]. In the case of HIV, it is believed that methylation of the 5′-LTR negatively regulates the binding of essential transcription factors, such as NF-κB and Sp-1 [62]. Demethylation of cytosine residue on the latent LTR resulted in activation of viral transcription of the HIV viral genome[63].

4. Promoter occlusion and convergent transcription

As mentioned earlier, HIV is capable of integrating into actively transcribed genes. But this also implies that latent proviruses could also integrate within highly expressed genes, suggesting that transcription of highly expressed cellular genes could interfere with viral transcription [64, 65]. Likewise, the establishment of latency also depends on how the viral genome integrates with the

host genome as different mechanisms of transcriptional interference may induce latency. When the proviral DNA has the same polarity as the host genome, promoter occlusion can occur. In this case, transcription of cellular genes by an elongating RNA Pol II interferes with assembly of the transcription factors essential for HIV transcription. The lack of proper assembly of transcription initiation complex results in HIV latency [66]. In contrast, when the proviral DNA has the opposite polarity compared to the hosting gene, RNA Pol II collisions may occur, resulting in premature termination of viral transcription. It was shown that orientation in the opposite direction as the host genome reduces HIV gene expression to more than tenfold [67]. Interestingly, a modest preference for integration with the same polarity as the host gene has also been reported [68].

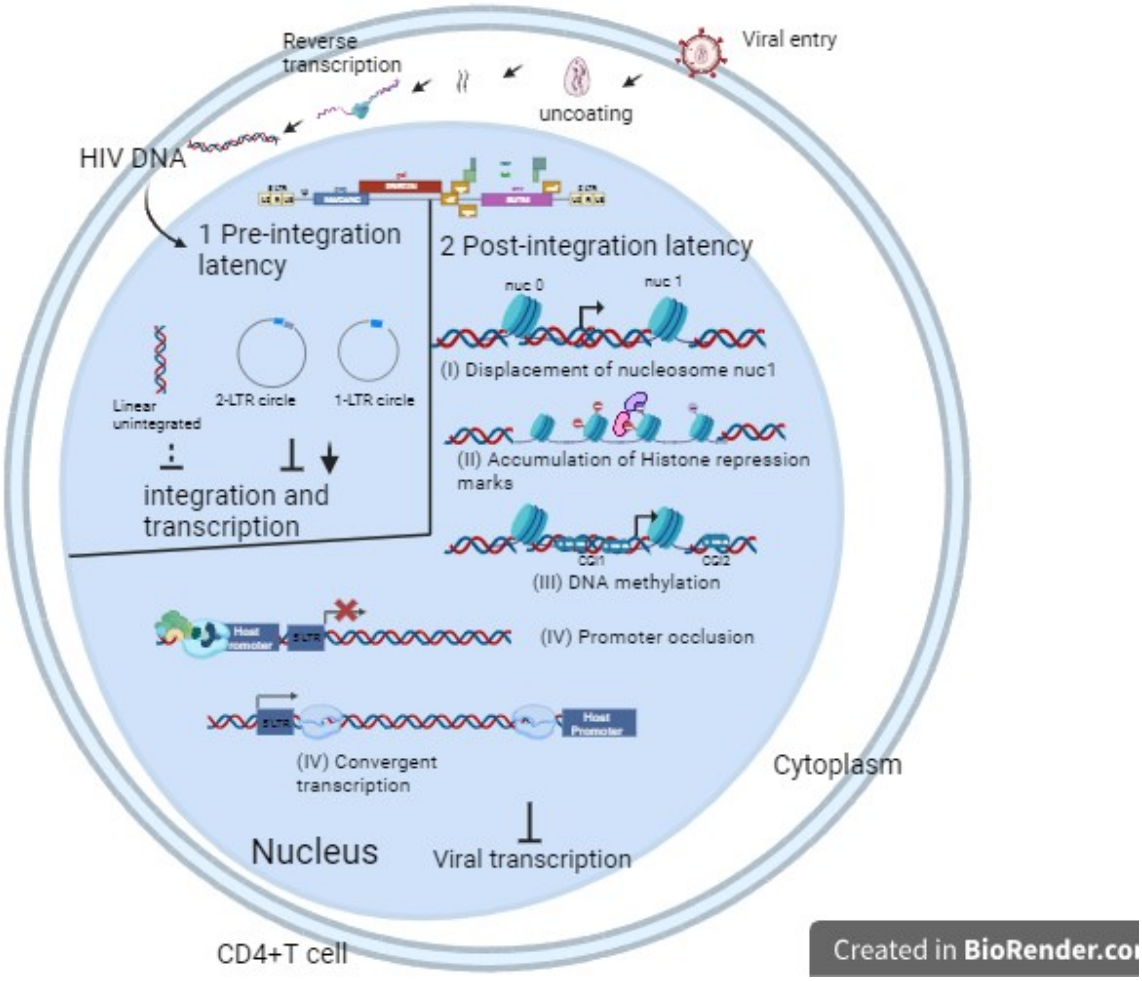


Figure 1.4- Epigenetic mechanisms leading to HIV latency
This figure shows various epigenetic mechanisms of HIV latency including pre-integration and post-integration latency mechanisms. This figure was made utilising Biorender.

1.4 Non epigenetic mechanisms

The **figure 1.5** shows the various non-epigenetic mechanisms resulting in HIV latency.

1.4.1 Ineffective tat transactivation / PTEFB regulation

Normally, RNA polymerase II (RNA Pol II) first binds to the promoters, initiates RNA synthesis, and then pauses in early transcriptional elongation to allow 5' capping of the newly formed messenger Ribonucleic acid (mRNA) before resuming the rest of the transcription. But, in the case of HIV transcription, due to the tightly bound nucleosome nuc-1 and negative regulators on RNA Pol II, the paused RNA Pol II is not able to resume its transcription directly. In this case, further signals are needed for the viral transcription to proceed from the paused RNA Pol II state to form a productive elongation complex [69]. This switch from promoter-proximal pausing to productive elongation is mediated by the coupling of Tat/ P-TEFb (positive transcription elongation factor b). P-TEFb is an essential elongation transcription factor constituted of two subunits: cyclin T1 (CycT1) and the cyclin-dependent kinase 9 (CDK9) [69].

Resting CD4+ T cells are characterized by extremely low levels of cycT1 due to actions of specific miRNAs and of the cellular factor NF-90, which blocks translation of CycT1 mRNA [70]. On top of that, in the absence of Tat, the P-TEFb complex is sequestered within the 7SK small nuclear ribonucleoprotein (snRNP) repressive complex, which is composed of the 7SK snRNA, the hexamethylene bisacetamide inducible protein 1 (HEXIM1), the 5'methylphosphate capping enzyme (MePCE), the La-related protein (LARP7), that are inhibited by the combined action of two repressors namely, negative elongation factor (NELF) and the 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor (DSIF). Even upon cellular activation, when Tat is not produced yet, P-TEFb is disassociated from the HEXIM-1/7SK snRNA complex but bound by the bromodomain protein 4 (BRD4), forming an active P-TEFb complex.

When Tat is able to compete with the BRD4 complex, then P-TEFb can be recruited to the HIV LTR [71].

Once sufficiently expressed, Tat is able to compete with BRD4 for binding P-TEFb, and Tat is also able to directly disrupt the inactive form of P-TEFb from the HEXIM-1/7SK snRNA complex. That way, Tat is able to recruit P-TEFb to the HIV promoter through TAR (Transactivation response element) and leading to transcription elongation [72, 73]. TAR is an RNA stem loop present in the R region of the HIV LTR. Tat can also recruit, in addition to P-TEFb, other elongation factors (such as ELL2, AFF4, ENL, and AF9), thereby forming the super-elongation complex (SEC). P-TEFb is thereby positioned to phosphorylate the C-terminal domain (CTD) of RNA Pol II, resulting in efficient elongation of viral transcription [74]. Thus P-TEFb, present in two forms including a free active form and an inactive form. The balance between these two forms controls the activity of P-TEFb and the subsequent viral transcriptional reactivation processes. Since, P-TEFb function is highly essential for HIV gene expression, the factors that alter functional P-TEFb are expected to favor the survival of latently infected cells [71].

1.4.2 The cytoplasmic localization of transcription factors in memory cells or cells entering memory phenotype

The transcription factor NF-κB plays a complex role during the replication of HIV. On one hand, NF-κB is crucial for induction of efficient proviral gene expression, and on the other, NF-κB activation also induces expression of genes involved in the innate immune response and the cellular antiviral response. NF-κB is usually sequestered in the cytoplasm of unstimulated cells as an inactive form by interacting with an inhibitory protein IκB. Following cellular activation, the phosphorylation of IκB by IKK (IκB kinase) leads to its ubiquitination and proteosomal degradation, allowing the translocation of NF-κB into the nucleus and the transcriptional trans-

activation of NF-κB-dependent genes. Notably, NF-κB also stimulates HIV transcriptional elongation by interacting with P-TEFb and directs the recruitment of a co-activator complex of HATs to the HIV LTR [75].

Additionally, the transcription factor Ap-1 binding site which is just upstream of the NF-κB element in the HIV 5'LTR is said to aid in establishing a latent infection [76], evidently, a deletion of this binding site was shown to strip HIV of its ability to establish a latent infection. This observation supports the idea that HIV latency is highly dependent on transcription factors [76]. The lack of active forms of other key cellular transcription factors (such as NFAT and STAT5) also leads to repression of viral transcription in resting CD4+ T cells [19]. Abdel-Mohsen et al. also showed that human Galectin-9 (Gal-9) is an effective mediator of HIV transcription and thereafter reactivation [77]. Recombinant Gal-9 potently reverses HIV latency in vitro and ex vivo through the induction of several HIV transcription factors (mainly NF-κB, Ap-1, and NFAT), expression and the inhibition of several chromatin modification and remodeling factors (including HDAC1, 2, and 3, EZH2, DNMT1) and ultimately gene expression [77].

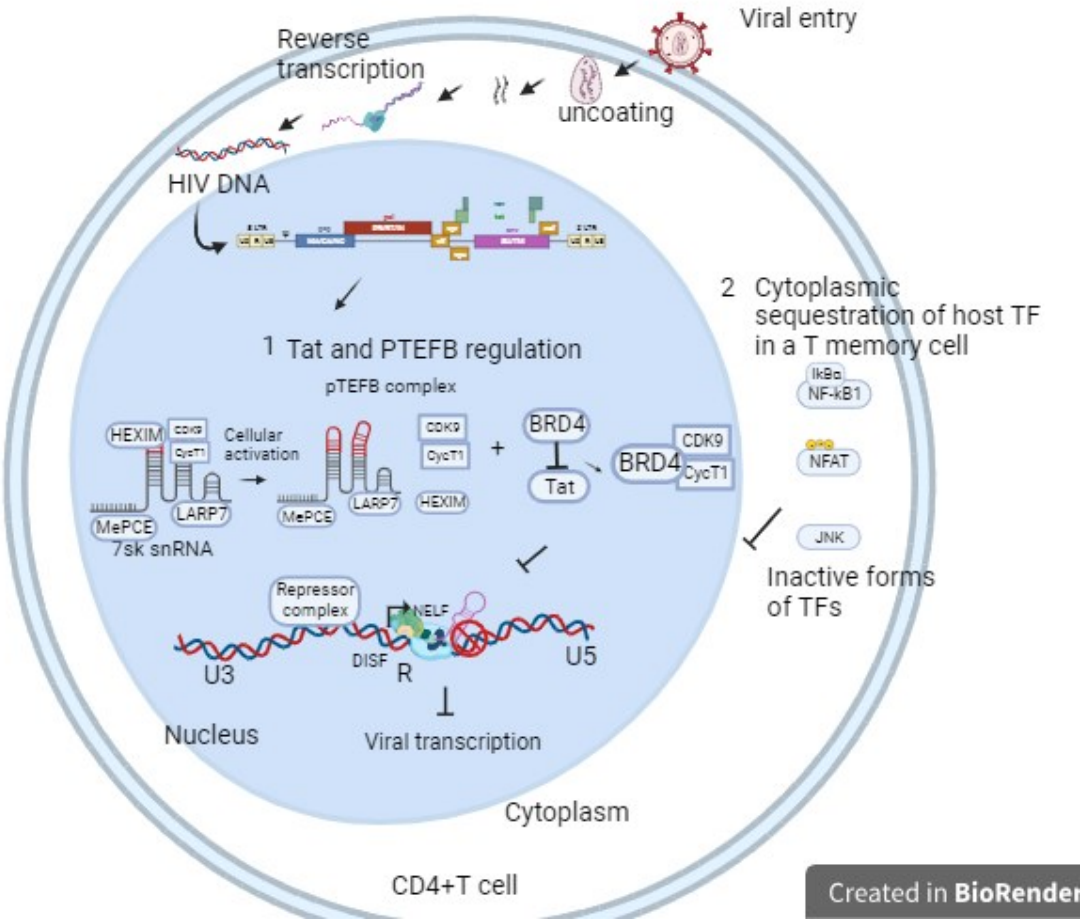


Figure 1.4- Non-Epigenetic mechanisms leading to HIV latency
This figure shows various non-epigenetic mechanisms of HIV 1 latency including Tat/ PTEFB regulation and cytoplasmic sequestration of host transcription factors. This figure was made utilising Biorender.

1.5 Viral molecular latency

As discussed in the previous sections, various mechanisms resulting in latency are, lack of active forms of cellular transcription factors that are essential for HIV gene expression in CD4+ memory T cell [12, 78] and HIV Tat and its cellular cofactors not being able to bind TAR [79] limits the initiation and elongation, respectively, of viral transcription [80, 81]. The latent reservoir may thus be established when a HIV infected activated CD4+ T cells reverts back to a memory state. In addition, DNA methylation and repressive histone modifications may silence proviruses [81]. Apart from the mechanisms mentioned above, recent evidence suggests that HIV latency can also be caused by mutations in the proviral genome [82].

Ho and group studied the latent reservoir and discovered that hypermutated proviruses are replication-defective due to numerous start codon mutations [27]. The hypermutated provirus is unable to produce functional viral proteins as a result of pre-mature stop codons introduced in most open reading frames (ORFs). From their viral outgrowth assay and subsequent sequencing, they demonstrated that a large fraction of non-induced proviruses is non-functional due to large internal deletions, assumed to be likely introduced during reverse transcription. Thus, any mutations present in cis elements (promoters, enhancers, and silencers) makes a perfectly intact provirus defective. They also found that most of the non-induced proviruses were integrated into active transcription units, consistent with previous studies that showed that in cell lines that most of the HIV proviruses were found integrated into introns of actively transcribed genes [65]. Intact proviruses undergo induction after cellular activation, some will be induced by one round of activation, while others will remain silent but retain the potential to be activated subsequently when conditions change [27]. This leads to inability to estimate the exact number of inducible viruses in the latent reservoir. For example, VOAs (viral outgrowth assays) could underestimate

the number of intact inducible proviruses because of delayed viral rebound exhibited by a few viruses, and at the same time, detection of defective proviruses by PCR based assays could overestimate the size of LR (latent reservoir) resulting in prolonged, excessive exposure to toxic latency reversing agents for a small population of LR. Thus, the molecular analysis of inducible proviruses could contribute to developing new HIV eradication strategies [83].

Comparison of reservoir measurements revealed that there is greater than 2-log discrepancy between the frequency of integrated proviruses measured by DNA PCR when compared to replication-competent proviruses measured by VOA showing a smaller number of replication-competent proviruses measured by the latter [84]. This discrepancy could be due to irregular frequency of defective proviral DNA and variable inductions of intact proviruses [27]. Most of integrated HIV proviral DNA are defective, hence measuring inducible viruses by VOA is more applicable than PCR based assays. The outgrowth assays are able to measure the replication-competent reservoir, however they do not capture all replication-competent proviruses.

Mutations arising in the viral promoter will result in interruptions in viral replication and gene expression because of loss of viral transcription [85]. We know that reverse transcript is an error prone process that ends up creation mutations in the viral genome. G-to-A mutations are the most frequent base pair mutations observed in HIV genome. Apart from reverse transcriptase, APOBEC3 has the ability to introduce such mutations in the viral genome. APOBEC3 [86] is a host produced retroviral restriction factor that is functionally different from other cellular restriction factors including TRIM5α [87, 88], tetherin [89], SAMHD1 [90, 91], Viperin [92], Mx2 [93], as APOBEC3 restricts HIV replication by introducing pre-mature stop codons in the viral genome.

1.6 APOBEC3

APOBEC3 (A3) belongs to the family of APOBECs. There are eleven members identified in this family, including APOBEC (A1), APOBEC2 (A2), APOBEC3A-H (A3A, A3B, A3C, A3DE, A3F, A3G, A3H), and APOBEC4 (A4), activation-induced cytidine deaminase (AID) [94].

The 7 Apolipoprotein B mRNA editing enzyme, catalytic protein-like 3 (APOBEC3) genes in the human genome code for APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H [95, 96]. Five of the seven members of A3 are A3G A3F, A3D, certain haplotypes of A3H, and one polymorphic variant of A3C have demonstrated antiviral activity against HIV [82]. A3s are polynucleotide cytidine deaminases, meaning they turn cytosine residues into uracil in negative sense single-stranded DNA that is freshly formed during reverse transcription. This is reflected as guanine to adenine (G-to-A) transition on the positive strand of HIV DNA [94, 97]. A3 proteins are effective against HIV, other Lentiviruses and Hepatitis B virus [98]. This catalytic mechanism of deamination observed in APOBEC3 (A3) was also found in *Escherichia coli* (*E. coli)* cytidine deaminase (showing functional similarity to AID and APOBEC1) [99]. This catalytic deamination could result in hypermutation of the viral genomic DNA potentially affecting its transcription and infectivity [97].

1.6.1 Deaminase-dependent mechanism of HIV restriction

Figure1.6-a shows how the A3 proteins are packaged along with the viral RNA genome inside the capsid of the egressing virus and how the viral protein VIF antagonizes A3 (**Figure1.6-b)**. When the A3 containing virus infects a cell, the A3s begin to mutate the HIV negative strand DNA during reverse transcription by inducing dC-to-dU deamination. This leads to introduction of G-to-A substitutions in the plus-strand DNA, resulting in missense and stop-codon mutations [100]. A3G and A3F are shown to restrict HIV more potently than the rest of the A3s [82]. A3G prefers to

deaminate 5'CC* sequences [101] whereas; A3F prefers to mutate 5'TC* [102]. This phenomenon of excess G-to-A mutations, first described in avian spleen necrosis virus (a gammaretrovirus) is referred to as G-to-A hypermutation [103, 104].

These A3 generated G-to-A hypermutations have the capability to restrict HIV replication, however HIV overcomes A3 restriction by producing the viral protein Vif, which targets and degrades A3, thereby rescuing HIV replication and infection. The viral protein Vif is produced only after the start of infection and increases it with time. These Vif proteins bind A3 proteins and then recruit E3 ubiquitin ligase complex for polyubiquitination and eventually degradation of A3 (as shown in **Figure 1.6-b**). When all A3s are degraded then the virus becomes fully infectious and is not further restricted by A3[105].

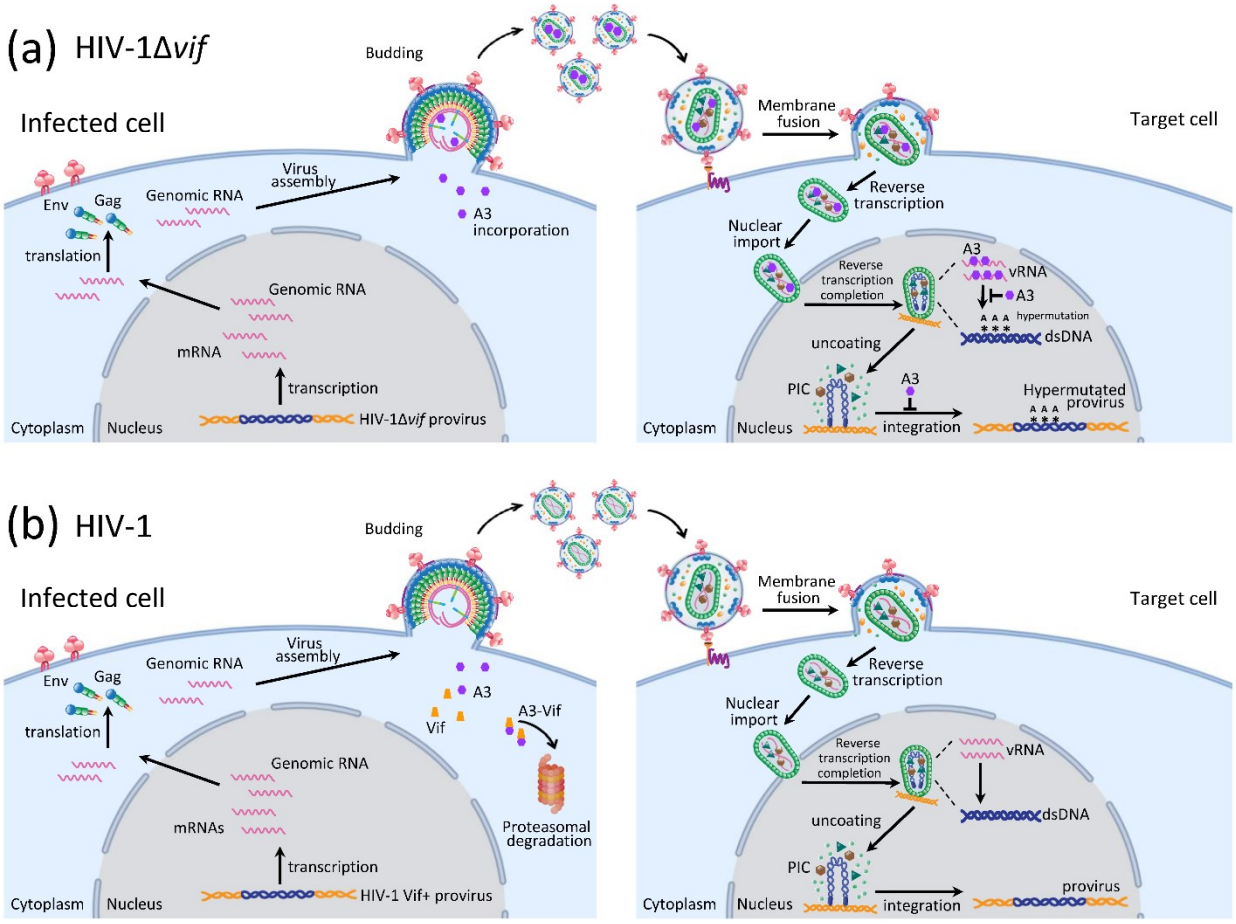
Figure 1.6 Interaction of APOBEC3 with the HIV VIF

A3-mediated HIV restriction and its counteraction by Vif.

(a) A3G inhibition of HIV Δvif. In the absence of a functional Vif protein (HIV Δvif), A3G is packaged into HIV virions in the producer cells, exerting its antiviral activity in the target cells by inducing lethal hypermutation.

(b) Vif induces degradation of A3 proteins. In the presence of wild-type HIV, expression of Vif in the producer cells targets A3 proteins for proteasomal degradation, leading to productive HIV infection of target cells.

Figure reproduced from reference [99] (open access)


(a) HIV-1Δvif
Budding
Infected cell
Virus assembly
Genomic RNA
Env
Gag
translation
A3 incorporation
Genomic RNA
mRNA
transcription
Cytoplasm
Nucleus
HIV-1Δvif provirus
Membrane fusion
Target cell
Reverse transcription
Nuclear import
Reverse transcription completion
A3
vRNA
A3
hypermutation
dsDNA
uncoating
PIC
A3
integration
Hypermutated provirus
(b) HIV-1
Budding
Infected cell
Virus assembly
Genomic RNA
Env
Gag
translation
A3
Vif
A3-Vif
Genomic RNA
mRNAs
Proteasomal degradation
transcription
Cytoplasm
Nucleus
HIV-1 Vif+ provirus
Membrane fusion
Target cell
Reverse transcription
Nuclear import
Reverse transcription completion
vRNA
dsDNA
uncoating
PIC
integration
provirus

1.6.2 Deaminase-independent mechanism of HIV restriction

A3 proteins can also restrict HIV replication via deaminase independent mechanisms, for example- A3 binding the viral RNA or viral enzymes, thereby reducing vDNA synthesis. The interaction of A3G and A3F with IN and RT were previously discussed, but their role in viral integration is not clearly resolved [106, 107]. As it was proved that both A3F and A3G proteins have the ability to compromise viral integration efficiency by inefficient modification of the long terminal repeats (LTR) of the virus [98], however it is still not known how this could potentially alter HIV proviral integration site selection.

1.7 Cure strategies

There are many strategies proposed for attaining HIV cure. These include approaches to either control viral replication without ongoing treatment (functional cure) or a complete elimination of infectious virus (sterilizing cure) from the patient. The former strategy aims at cure wherein the patient is able to lead a fully functional life without the need for ART, it encompasses the either the use of therapeutic vaccines to help promoting HIV specific immune responses allowing immune-mediated suppression of viraemia or by using small molecule inhibitors to completely silence the viral transcription. By contrast, the latter strategy involves a significant reduction in the size of the latent reservoir by reactivating the viral transcription so that immune response can clear infected cells to prevent viral rebound. [18]

1.7.1 Shock and kill strategy

Since the latent reservoir is not transcriptionally active, no HIV specific gene expression is detected [108]. The shock and kill strategy is a sterilizing cure strategy that makes use of latency-reversing agents (LRAs) to induce HIV gene expression, allowing for infected cells to be detected and cleared off by cytolytic immune cells [108]. Several LRAs have been tested in *in-vitro* and *in-vivo* studies [109], but none of them are yet approved for use as a LRA. A possible HIV eradicating strategy should involve using LRAs in patients on ART [110]. So that, the cells harboring induced proviruses could then be lysed by HIV-specific cytolytic T lymphocytes (CTL) [111], and new rounds of infection can be blocked off by ART.

The drugs categorized as LRAs are encompassed in one of the several classes mentioned below. They are protein kinase C agonists [112, 113], histone deacetylase inhibitors (HDACi) [87, 111], TLR agonists [114], PTEN agonist [115] etc. Most of these drugs are approved for treating some type of cancer but are not approved for use as a latency reversing agent as mentioned in the

previous section. Although these drugs show some extent of HIV induction when tested *in-vitro*, they unfortunately failed to be approved for human consumption as an LRA in clinical trials.

The LRAs tested *in-vivo* so far have mainly focused on reversing epigenetic blocks that prevent HIV gene expression using histone deacetylase inhibitors (HDACi) such as vorinostat, romidepsin and Panobinostat [116, 117]. Vorinostat (Vor), which is also called Suberoylanilide hydroxamic acid (SAHA) was used in clinical studies to disrupt HIV latency and activate viral expression in HIV-infected patients. A study that involved treatment with a single dose of Vor, increased the cellular acetylation biomarkers, and also led to an increase total HIV RNA expression in resting CD4+ cells [116]. But it does not help in reducing the size of latent reservoir [118].

Protein kinase c agonists are bryostatin, PMA and prostratin. They activate cellular isoforms of PKC thereby initiating NF-κB signalling to induce HIV transcription potentially leading to latency reversal. [119]. The host transcription factors essential for HIV gene expression are either absent or sequestered as inactive forms in a resting CD4+ T cell, so these cells have to be activated to allow nuclear import of these transcription factors (NF-κB and NFAT) [13, 120, 121].

Overall, PMA/I is regarded as a global T cell activator, and most of the LRA candidates were ineffective at reversing latency ex vivo in cells collected from people living with HIV (PLWH) and, so far, they have failed to shrink the latent reservoir *in-vivo* [118, 122, 123].

1.7.2 Block and lock strategy

The block-and-lock approach aiming for a functional cure, makes use of epigenetic silencing of the proviral genome thereby locking the virus in a permanently latent state, unable to be reactivated [124].

Block and lock is a recently developed strategy to silence HIV transcription so that people living with HIV don't require life long ART. Ideal drug candidates that promote transcriptional silencing,

in addition to favorable pharmacokinetics and minimal toxicity, should “block” transcription initiation and/or elongation resulting in a durable “lock,” or epigenetic silencing of the promoter, such that little or no residual viral production occurs in the absence of ART and the immune system holds the virus in check. An approach to silence HIV transcription could be achieved either by targeting viral factors or host factors. Targeting viral factors for transcription inhibition can be achieved by targeting Tat-TAR interaction as this is a rate-limiting step [125]. In addition, targeting host factors including the most common targets- positive elongation factor (P-TEFb) [126, 127], Xeroderma pigmentosum Type B (XPB) (XPB is an important subunit of the TFIIH complex) [128], Heat-Shock Protein 90 (Hsp90) (found to directly associate with the HIV promoter) [129], Bromodomain-Containing Protein 4 (BRD4) [130, 131] and many more could be used to inhibit and lock the latent virus.

Some of the approaches currently being explored to mimic such a functional cure include modulation of HIV transcription [132, 133], therapeutic vaccination to induce effective HIV-specific CD4+ and CD8+T-cell responses [134] and long-term expression of broadly neutralizing antibodies or entry inhibitors with gene therapy vectors [135].

RATIONALE

Among the many factors responsible for viral latency, we are interested in viral molecular latency. At the molecular level, one of the most significant mechanisms of latency is a loss of viral transcription which leads to an absence of virus production. By examining viruses produced from the latent reservoir, most are defective having G-to-A hypermutations introduced within an APOBEC3 genetic context, but some appear intact and potentially capable of reseeding the latent reservoir. VIF-mediated degradation of cellular APOBEC3 leads to fewer molecules of APOBEC3 being introduced into egressing viruses , thereby leading to gradually lower levels of sub-lethal G-to-A mutations in the viral genome that is expected to lead to tendencies of a latency-like phenotype.

HYPOTHESIS

We hypothesize that low levels of "G-to-A" transition mutations in the HIV LTR region introduced by APOBEC3G/F lead to a latency-like phenotype.

2 OBJECTIVES

1. Analyze the effects of A3 mutations on promoter activity.
2. Evaluate the infection, induction to LRAs, and integration levels of the LPVs.
3. Determine how the LPVs evolve over a period following induction.

3 MATERIALS AND METHODS

3.1 Tissue Culture Cell Propagation

For culturing HEK 293T cells (Human Embryonic Kidney Epithelium 293T cells) (ATCC # CRL-3216), Dulbecco's Modification Eagle's Medium (DMEM) (Wisent, #319-005-CL, St Bruno, QC, Canada) were supplemented with 100U/mL penicillin and 100ug/mL streptomycin (P/S) (GE Healthcare, Chicago, IL, USA) and 10% v/v Fetal Bovine Serum (FBS) (Corning 35-015-CV), hereafter referred to as complete DMEM. HEK 293T cells were passaged every 72 hours approximately.

For culturing Jurkat cells (ATCC #TIB-152), RPMI 1640 (Wisent, St. Bruno, QC, Canada) were supplemented with 100U/mL penicillin and 100ug/mL streptomycin (P/S) (GE Healthcare, Chicago, IL, USA), 10% v/v Fetal Bovine Serum (FBS) (Corning 35-015-CV), hereafter referred to as complete RPMI. Cells were passaged every 72 hours.

For culturing primary human CD4+ T cells, complete RPMI was used. In addition, the CD4+ T cells were activated after isolation by supplementing the complete RPMI with 2.5ug of PHA and 20 IU of Interleukin 2 (Il2) and incubating the cells at 37 ºC for 72h.

For culturing U87 CD4+ CXCR4+ (NIH AIDS #4036) (hereafter referred to as U87 cells), complete DMEM was supplemented with 300ug/mL G418 and 1ug/mL puromycin. U87 cells were passaged every 4 to 5 days.

For culturing Ghost CD4+ CXCR4+ CCR5+ (NIH AIDS #3942) (herein referred to as Ghost cells), complete DMEM was supplemented with 500ug/mL G418, 100ug/mL hygromycin and 1ug/mL puromycin. Ghost cells were passaged every 3 to 4 days.

All cell lines mentioned in this thesis were cultured in a 37 ºC + 5% CO2 incubator (Forma Series II Water Jacket CO2 Incubator) (Thermo Scientific Inc., Burlington, Ontario, Canada).

3.2 Latency-Prone Virus Plasmid Creation

pEGFP-C3 plasmids and non-replicative pNL4.3-Δenv-Δvif plasmids harbouring the various LTR mutation combinations were previously created in the lab by Cindy Lam (refer to appendix 8.1). These were re-cloned in the lab recently to eliminate certain extra unwanted mutations detected during sequencing by Dr. Samar Dankar. The replicative CXCR4 tropic pNL4.3 plasmids were cloned previously by Mathew Greig (refer to appendix 8.2) using the non-replicative plasmids mentioned above.

3.3 Plasmids used for virus production

pEGFP-C3 backbone + mutated LTR – has the mutated LTR as a promoter to express GFP. pNL4-3ΔenvΔvif backbone + mutated LTR- non-replicative allowing only one round of infection with co-transfected with VSV-G. The mutants are called non-replicative LPVs (Latency-prone viruses) CXCR4-tropic pNL4 backbone + mutated LTR- Fully replicative, the mutants are called replicative LPVs. VSV-G – Pseudo typing plasmid to help viral entry of the non-replicative LPVs.

3.4 Plasmid Stock Propagation

To propagate LPV plasmid stocks to be used in transfections, along with a vesicular stomatitis virus glycoprotein (VSV-G) plasmid (Addgene, Cambridge, Massachusetts, USA), 2ul of each plasmid was mixed well with 50uL of Stbl-2 *E. coli* cells. This mixture was then incubated on ice for 30 minutes. Tubes were then mixed gently by tapping motion and were then subjected to heat

shock in a 42 ºC water bath for about 90 seconds, followed by 2 minutes of incubation again on ice. Each transformant was taken in separate Eppendorf tubes and 1mL of 2XTY broth (16g bacto tryptone, 10g bacto yeast extract, 10g NaCl, 1L ddH2O, pH 7.2) was added to it, and then incubated for 90 minutes in a 600rpm shaker at a temperature 30 ºC. Then, 200uL of the bacterial broth that was newly transformed was spread using an L shaped spreader onto LB + Amp or Kan plate (20g LB powder, 17g bacto agar, 1L ddH2O, 100mg ampicillin/ Kanamycin) (prewarmed) according to the plasmid propagated that was prewarmed. Plates were incubated in a 30 ºC incubator overnight or for 20-24h, post-colony formation, one colony per transformation was selected and picked with a regular 200uL pipette tip and the tip was ejected into 100uL of LB medium and ampicillin/Kanamycin containing Erlenmeyer flask/ 3mL of LB medium in a 5mL bacterial culture flask. Flasks were incubated overnight in a 30 ºC shaker. Midipreps/Minipreps of all the 3'LTR or LPV plasmids were performed by utilizing an E.Z.N.A Plasmid DNA Midi Kit (Omega Biotek Inc. #D6904-04, Norcross, GA, USA) eluting the plasmid DNA in 700uL/ 50ul of sterile RNAse-free water. The extracted plasmid DNA were stored at -20 ºC freezer as stocks and working solutions of each plasmid was diluted to 100ng/uL in sterile RNAse-free water and stored at 4 ºC.

3.5 Assessment of functionality of the mutated promoters

24h prior transfection, HEK 293T cells were seeded at a density of 1.0 x 10^5 cells per well of a 24-well plate in 0.5 mL of complete DMEM. The cells were transfected with 200 ng of each clone in a peGFP backbone using GeneJuice transfection reagent. The transfection was followed by the addition of 10ng/ml of PMA and 1uM of Ionomycin or DMSO. 24 hours post-transfection and PMA/I treatment, the PMA/I in the media was substituted with regular DMEM. 24 hours after the

media change, cells were collected, and the promoter functionality of the mutated LTRs was assessed by flow cytometry measurements of the eGFP expression and fluorescence intensity.

Similarly, 24h prior to transfection, HEK 293T cells were seeded at a density of 1.0 x 10^5 cells per well of a 24-well plate in 0.5 mL of complete DMEM. These cells were then co-transfected with 100 ng of each 3'LTR mutant and either 100 ng of the Tat expression plasmid or pcDNA (negative control) using GeneJuice transfection reagent. The cells were harvested 48 hours post-transfection, and the promoter functionality of the mutated LTRs was assessed by measuring the eGFP expression and fluorescence intensity measured on FITC channel (530/30) using BD FACSCelesta and BD FACSDiva (Software v8.0.1)). Post-acquisition analysis was performed using FlowJo (software v10.4.2).

3.6 In-silico analysis of TFB efficiency of the mutated promoter

In order to check if the mutations of interest lead to compromised transcription factor binding, we have used an online tool "Fabian variant" developed by Robin Steinhaus et al. The transcription factors considered in our project are NK-κB, SP-1, NFAT, Coup, Ap-1, ETS-1 and CTF/NFI. The sequences obtained from sequencing was first copied from the notepad document and passed on Microsoft word. The extra sequences apart from the 634bp of the LTR were deleted in all the clones. Then, the consensus sequence of every transcription factor was compared to the wildtype sequence of HIV LTR. Finally, the mutated TFB sites of each clone were compared to the wildtype sequence for every transcription factor mentioned above to check if there is any difference in the binding affinities between the 2 sequences compared. The output obtained from the software for every clone is a number between -1 to +1. This output was used to create a heat map on Graphpad prism.

3.7 Viral Transfections

HEK 293T cells were prepared for transfection 24h prior. The cells were centrifuged at 500xg for 5 minutes, resuspended in complete DMEM and then the cell count was determined using a hemocytometer, and the cells were plated in 6-well tissue culture plates with a density of 625,000 cells per well and a final volume of 2mL complete DMEM per well (VWR International #10062-892, Mississauga, ON, Canada).

On the day of transfection, 1.7mL Eppendorf tubes were used for each transfection setup. The appropriate amounts of VSV-G plasmid (250ng), LPV plasmid (750ng) are required per well in the case of a 6 well plate. For every 1000ng of plasmid DNA, 3ul of genejuice and 100ul of serum-free DMEM were added as a master mix and mixed by vortexing following which it was incubated for 5 minutes in room temperature. Appropriate volume of the mixture was then added to each plasmid tube, tapped 8 times to mix and incubated at room temperature for 10 minutes before adding to the cells. After the 10-minute incubation, the tubes were mixed 3 times by inversion, and then each transfection solution was added dropwise to the respective wells. Plates were mixed by gentle biaxial shaking and placed in a 37 ºC + 5% CO2 incubator for 72hr. The replicative viral transfections were performed in the CL2+ facility of our lab.

After 72 hours, the supernatant containing the virus was collected from each well in separate 5mL Falcon polystyrene, FACS tubes (Corning #352054, Tewksbury, Massachusetts, USA). The FACS tubes containing the virus were centrifuged at 500xg for 5 minutes and filtered through a 0.45um filter (VWR #CA28143-312, Canada) to remove cellular debris. The supernatants were then stored at 4 °C. To pick up the cells, 500uL of 1X PBS with 5mM EDTA was used. The cells were centrifuged at 500xg for 5 minutes followed by a 1mL 1xPBS wash and centrifuged again. The cells were then resuspended in 1.2mL of basic sort buffer (1X PBS, 5mM EDTA, 25mM Hepes,

1% globulin-free bovine serum albumin, pH 7.2-7.4, 0.2um filtered) followed by fixing with the addition of 1.2mL of 4% paraformaldehyde (PFA). The transfection efficiency of viruses was assessed by measuring the eGFP expression and fluorescence intensity measured on FITC channel (530/30) using BD FACSCelesta and BD FACSDiva (Software v8.0.1)). Post-acquisition analysis was performed using FlowJo (software v10.4.2).

3.8 Viral titration with ELISA p24 (gag) Quantification

Previous researchers in the lab prepared anti-p24 antibodies 183-H12-5C and (biotinylated) 31-90-25 for use in a p24 sandwich ELISA to measure the amount of virus in each viral transfection. 100ul of anti-p24 183-H12-5C coating antibody at a concentration of 5ug/mL was used to coat high binding 96-well ELISA plates (Santa Cruz Biotechnology #204463, Dallas, TX, USA) (refer to appendix 8.4) in 0.1M sodium bicarbonate buffer (8.4g $NaHCO_3$ in 1L sterile Mili-Q water, pH 9.5). The plates were incubated on a slow rotator at 4 °C for 2 or 3 days. Afterward, the plates were washed thrice with ELISA wash buffer (1X PBS + 0.05% Tween-20) using the plate washer. Then, 200ul of ELISA blocking buffer (ELISA wash buffer, 5% milk, 2.5% globulin-free bovine serum albumin) was added to each well for blocking at room temperature for 2 hours on a shaker. Recombinant HIV p24 standard (Abcam #ab43037, Cambridge, MA, USA) was diluted in DMSO and divided into aliquots starting at 200ug/mL and further diluted to 100ug/mL and 10ug/mL. The 10ug/mL aliquots were used for the ELISAs. The standard p24 was then diluted from a concentration of 10ug/mL to 100ng/mL, which is a 1:100 dilution (10uL of 10ug/mL p24 + 990uL ELISA blocking buffer) in ELISA blocking buffer, with 7 serial dilutions of 1:2 down to 0.78125ng/mL.

The virus supernatants were warmed in a 37 °C incubator for 10 minutes and then diluted starting with 1:100 in ELISA blocking buffer (10uL of viral supernatant + 990uL ELISA blocking buffer). Serial dilutions of 1:2 was performed down to 1:200, 1:400, 1:800, 1:1600 and 1:3200 depending on the LPV. Each sample was mixed by vortexing at each step of the dilution process. ELISA lysis buffer (ELISA wash buffer, 2.5% Triton X-100, 0.02% Thimerosal) was added after ensuring each sample to have a final volume of 500uL, the tubes were mixed by gently inverting 8 times and then incubated for 30 minutes at 37 °C. The samples were then added in triplicates (125uL each) to the ELISA plates, which were washed and incubated on a rotator overnight at 4 °C.

The plates were washed again thrice, and 100ul of biotinylated p24 antibody 31-90-25 at a concentration of 840ng/mL diluted in ELISA blocking buffer was added to each well and plate was incubated for 2 hours on a rotator at room temperature. After another wash, 100ul of 1:10,000 diluted Streptavidin-HRP (ThermoFisher Scientific #S911, Waltham, MA, USA) in ELISA blocking buffer was added and incubated for 1 hour on a rotator at room temperature. The plates were washed 4 times, and SIGMAFAST OPD tablets were dissolved in 20ml of sterile water to be used as ELISA substrate. 100ul of the sterile water containing substrate was added to each well. The plates were covered with foil to conceal them from light and after 13 to 18 minutes, their optical densities were quantified with the help of BioTek ELISA plate reader at a wavelength of 405nm and 600nm.

The concentration in ng/mL of each LPV supernatant was calculated using Microsoft Excel 2016 by fitting a regression line that compared the quantified optical density to the fixed concentration of the standard curve. Regression lines with R^2 values above 0.9 were considered valid. The

equation of each regression line was then used to quantify all samples on the plate, and normalized infection volumes were calculated based on the estimated concentrations.

3.9 HEK 293T PMA/I induction Assay

HEK 293T cells were seeded at a density of 125,000 cells per well in 12-well plates (VWR International #10062-894, Mississauga, ON, Canada) with a final volume of 1mL complete DMEM, 24 hours prior to infection.

After 24 hours, the non-replicative viral supernatants produced from pNL4.3-Δenv-Δvif were centrifuged at 500xg for 10 minutes at 25 °C. The volume of virus to be added was removed from each infection well, and polybrene (Sigma-Aldrich, Oakville, Ontario, Canada) was added to a final concentration of 8ug/mL. The plates were gently mixed by biaxial shaking and incubated for 10 minutes at 37°C with 5% CO2. Each well received 345ng of LPV supernatant, and the plates were gently mixed again. A 1-hour spinoculation step at 900xg and 25 °C was performed, followed by placement of the plates in a 37 °C incubator with 5% CO2.

After 24 hours of infection, the medium was aspirated from each well and refreshed with 1mL of complete DMEM containing PMA and Ionomycin or a DMSO control (matched in volume to PMA/I). The plates were returned to the 37 °C incubator with 5% CO2. After 24 hours of activation, the medium was aspirated, and the cells were detached from the plates using 500uL of 1X PBS with 5mM EDTA. The cells were centrifuged, washed with 1mL of 1X PBS, and then fixed by adding equal volumes of basic sort buffer and 4% paraformaldehyde (PFA). Lastly, the tube was gently vortexed.

Analysis was performed on a BD FACSCelesta flow cytometer using BD FACSDiva software. The eGFP fluorescence was measured in the FITC channel (530/30). Analysis of the cells was performed using FlowJo (software v10.4.2) software.

3.10 Jurkat PMA/I induction Assays

Jurkat cells were prepared for infection by seeding 250,000 cells per well in 12 well plates in 500uL of complete RPMI, 24hours prior to infection. On the day of infection, the cells were topped up with more complete RPMI following the equation:

Volume of RPMI to be added = 1mL final volume required – 500uL old media – volume of 345ng of LPVs (according to ELISA results).

After topping up with RPMI, 1ul of 1000X polybrene was added, mixed gently by biaxial shaking, and incubated for 10 minutes at 37 °C with 5% CO2. 345ng of LPV supernatant was added to appropriate wells, and the plates were gently mixed again. A 1-hour spinoculation step at 900xg and 25 °C was performed, followed by placement of the plates in a 37 °C incubator with 5% CO2.

After 24 hours of infection, PMA 10ng/mL and Ionomycin 1uM were diluted in complete RPMI to 2 times the required concentration (i.e., PMA 20ng/mL and I 2uM). The plate was checked carefully to locate the Jurkat cell clumps. Next, 500uL of old media was carefully aspirated from the well leaving the cell clump intact. 500uL of the prepared complete RPMI + 2X PMA/I solution was then added each well. The plates were returned to the 37 °C incubator with 5% CO2. The Jurkat cell suspension in all the wells were harvested 24 hours post-activation, centrifuged at 500xg for 5 minutes and then washed with 1mL of 1X PBS. The cells centrifuged again and resuspended in 300uL basic sort buffer and fixed with 300uL 4% PFA and a gentle vortex in the end.

Analysis was performed on a BD FACSCelesta flow cytometer using BD FACSDiva software. The eGFP fluorescence was measured in the FITC channel (530/30). Post-acquisition analysis was performed using FlowJo software (software v10.4.2).

3.11 Integration

For determining the integration levels of each of the non-replicative LPVs having pNL4.3-Δenv-Δvif backbone, the Jurkat cells were infected and treated with PMA/I as mentioned previously. 24 hours post-activation, the cells were harvested as mentioned previously and while in 1xPBS, two-third of these cells were separated to extract genomic DNA. The remaining one-third of the cells were analysed by flow cytometry to measure the activation and infection percentages. The 2/3 cells were first centrifuged at 17,000xg for 1 min and 1X PBS aspirated off leaving the pellet intact. The extraction of gDNA from these cells was done using Wizard gDNA Purification Kit (Promega #A1120). Then the extracted gDNA was eluted in 100uL of gDNA rehydration buffer at room temperature overnight. The concentration of the extracted gDNA was measured using nanodrop and all the LPVs were diluted to a working conc of 10ng/mL in RNAse/DNAse sterile water. A two stage semi-nested Alu-qPCR was performed to measure the integration levels of all the non-replicative HIV LPVs in Jurkat cells. As the name suggests, the first stage was an Alu PCR amplifying between each Alu region in the cellular genome and U5 region of the integrated HIV LPV's LTR. The Alu PCR was performed with the following reactants: 50ng gDNA, 2uL Primestar GXL, 10uL 5X Primestar GXL Buffer, 4uL dNTPs, 26uL sterile water, 0.2uM Alu1 primer, 0.2uM Alu2 primer and 0.2uM Lambda R U5 Rev 1 primer (refer Appendix 8.3 for primer and probe sequences). Reactants were added to a semi-skirted 96 well qPCR plate and centrifuged at 300xg for 2min to let the liquid settle at the bottom. PCR reactions were performed in an

Eppendorf Vapo Protect Pro S Mastercycler. The reaction setting for Alu PCR were: 94 ºC for 1min; followed by 30 cycles of 98 ºC for 10sec, 55 ºC for 15 sec and 68 ºC for 10min; followed by a final 68 ºC for 10min and 10 ºC hold. Amplicons from the first run were subjected to 1/10 dilution in DNAse/RNAse free sterile water. The second run was either an LTR qPCR or Actin qPCR. LTR qPCRs were performed by taking 5uL of 1/10 diluted amplicons from the first run, 10uL of 2X PowerUp SYBR Green Master Mix (Thermo Fischer Sci. Ref# 45 A25742), 0.2uM lambda R Rev 2 primer, 0.2uM late U3 Fwd primer, 50nM LTR Probe and 3.75uL sterile water per reaction. On the other hand, Actin qPCRs were done with 5uL of LPV gDNA, 10uL of 2X PowerUp SYBR Green Master Mix, 0.2uM Actin 1 primer, 0.2uM Actin 2 primer, 50nM Actin Probe and 3.75uL sterile water per reaction. These reactants were taken in a qPCR 96 well plate and then quickly spun for 2 minutes at 300xg. The plates were sealed with an adhesive plate cover that is specifically designed for qPCR. LTR qPCRs were done using the below mentioned settings: 95 ºC for 3min; 60 cycles of 95 ºC for 15sec, 60 ºC for 30sec and 72 ºC for 30sec; followed by a melt curve of 95 ºC for 15sec, 60 ºC for 1min and 95 ºC for 15 sec. The CT values were first normalized using the actin CT values which were regarded as loading controls. Secondly, the integration level of the unactivated pNL4 wildtype was used to compare all the LPVs. The mean integration of the uninduced pNL4 wildtype was set to 1.0, and all other LPVs were normalized by this value.

3.12 CD4+ T cell infection and Activation Assays

CD4+ T cells were isolated from human PBMCs (Peripheral blood mononuclear cells). Blood was diluted 1:2 in 1X PBS+2mM EDTA buffer. Firstly, PBMCs were extracted from whole blood using the SEPMATE tubes (Stemcell technologies Catalog # 85450) and Lymphoprep (Stemcell

technologies Catalog # 07851) by gradient separation method. The PBMCs obtained at this stage were washed multiple times by subsequent centrifuging for 10 minutes at 300xg. Human CD4 isolation kit (Miltenyi biotec cat# 130-096-533) was used for negatively select CD4+ T cells from the mononuclear cells. When passed through a magnetic column, the labelled non CD4 cells retain in the column and the CD4+ T cells pass through. The cells before and after column were collected to check the purity by staining with CD3 and CD4 antibodies (BD biosciences cat# 560835, Biolegend cat# 317410). After counting, the cells were seeded 1000000 cells well with 500ul of complete RPMI substituted with 2.5ug of PHA (invitrogen) and 20 IU of Il2 (R&D systems cat# 202-IL) and one well of cells only in complete RPMI (unactivated control). This substitution helps in activation of the CD4+ T cells.

72h post-activation, the activated cells were collected in a 50ml falcon tube and centrifuged for 8 minutes at 300xg. The supernatant was discarded and resuspended in 25-35ml complete RPMI (decided based on the size of the pellet). The cells were then counted and seeded in a 48 well plate, each well having 500,000 cells in 100ul of media.

One well of the activated cells and the unactivated control were stained with CD3, CD4 and CD69 (Biolegend cat# 310904) in order to check if the T cells obtained are activated.

Replicative viruses were generated from LPV plasmids having a replicative CXCR4 tropic backbone by the viral transfection procedure previously explained. The viral titre was also determined using the p24 ELISA procedure mentioned above.

The cultured CD4+ T cells in 100ul of complete RPMI were infected with the LPVs equivalent to 2000ng of p24 calculated by elisa. The volume of media added to each well was obtained from the calculation:

Volume of media = 700-100-volume of virus to be added

While adding virus to each well, the cells and viral supernatants were mixed well by pipetting. The plate was then spinoculated for 2h at 1200xg at 25 ºC. The plate was returned to 37 ºC incubator after the spin for the next 24h. 24h after infection, the cells were transferred to separate FACS tubes, centrifuged for 8 minutes at 300xg and resuspended with complete RPMI+ 20IU Il2 + 0.1ng/mL PMA and 1uM Ionomycin (PMA/I activation). The cells from FACS tubes were transferred to fresh 48 well plate and returned to the 37 ºC incubator. 24h after PMA/I activation, the cells were transferred to FACS tubes, centrifuged at 300xg for 8 minutes and washed with 1XPBS. The cells were centrifuged again and resuspended with 500ul of complete RPMI+ 20IU of Il2. The cells were re-transferred to a fresh 48 well plate and returned to 37 ºC incubator.

48h post media change and 72h post-activation, the cells were collected in separate FACS tubes, spun for 8 minutes at 300xg for washing. The cells were re-suspended in 100ul of 1xPBS per tube and stained with the following fluorophores, 2.5ul CD4-PE, 2.5ul of CD3-PerCP-Cy5.5 and 2.5ul of CD45-Pacific blue (Biolegend cat# 304029) and mixed well. The cells were incubated at 4 ºC for 30 minutes to allow staining. The cells were washed with PBS as mentioned previously. Half of the cells were permeabilized using 200ul of the Cytofix perm solution from the perm kit (BD biosciences cat# 554714) and the other half was fixed with 125ul each of BSB and 4%PFA. The cells were fixed/permeabilized for 20 minutes in the fridge. Post-permeabilization, the cells were washed twice with the Cytowash solution that was provided with the perm kit. The PFA fixed cells

were just washed with 1xPBS instead. After the second wash, p24-APC antibody (MediMabs cat# MM-0289-APC) was diluted in cytowash to add to permeabilized cells to internally observe p24 expression. The final concentration of the anti-p24 antibody is 1.25ug/mL. The PFA fixed cells were also stained with p24 APC which was diluted in 1XPBS to observe external p24 expression. The cells were allowed to stain in the fridge for 30 minutes. Post-staining the cells were washed twice again and finally resuspended with 400ul of 1XPBS. All the infected and stained samples along with single color controls were analyzed on BD Fortessa using channels PE (585/20) APC (660/20) Pacific blue (450/50) PerCP Cy5.5 (695/40) and FITC (530/30), and post-acquisition analysis with Flow Jo software to obtain the percentage of infection.

3.13 Evolution experiments

The viral transfection was performed as previously mentioned to generate replicative LPVs and the wildtype pNL4 control. The viral titre was calculated using the above ELISA p24 quantification protocol from the transfection supernatant. **Figure 3.1** is a graphical representation of the protocol used for this experiment. On day 0, 6 hours before infection, 100,000 U87 cells were cultured in a 12 well plate in 500uL of complete DMEM (without antibiotics). During infection, 0.5uL of 1000X Polybrene was added to each well and the cell were then mixed by gentle biaxial shaking for about 30 seconds and incubated at 37 ºC for 10min. Meanwhile, the viral transfection supernatants were also incubated at 37 ºC to warm them up. Virus volume equivalent to 25ng of p24 of each LPV supernatant was then pipetted in dropwise to each well, and then the cells were again mixed by gentle biaxial shaking. After mixing, plates were spinoculated at 900xg for 1hr at 24 ºC and were then placed in 37 ºC +5% CO2 incubator. 24 hours post-infection, 500ul of DMEM+ 10ng/mL PMA + 1uM Ionomycin was added making the final concentration of PMA

at 5ng/mL and 0.5uM Ionomycin. 24 hours post-activation, media containing PMA/I was aspirated and replaced with 1.5mL of complete DMEM after a 1mL 1X PBS wash before placing it back in the incubator.

7 days after infection and 6 days post-activation, virus media from each well was collected, centrifuged at 500xg x 5min and then filtered through 0.45um filter in order to remove cellular debris and finally stored at 4 ºC until using for further infections. The cells in the wells were gently washed with 1X PBS and then 1.5mL of complete DMEM without PMA/I was added to each well before returning the plates back in to the 37 ºC incubator for the next 7 days. Ghost cells were seeded in a 24 well plate, with each well having 12,500 cells in 500uL of complete DMEM without antibiotics on the same day. 24 hours after preparing Ghost cells, 0.75uL of 1000X polybrene was added to each well, and then mixed by gentle biaxial shaking, and incubated for 10 minutes at 37 ºC. Meanwhile, filtered supernatants from day 7 that were stored in the 4 ºC fridge were warmed up in 37 ºC incubator. After 10 min, 250uL of day 7 LPV supernatant was added to each well having Ghost cells, plates were mixed by biaxial shaking and then spinoculated at 900xg for 1 hr at 24 ºC before placing the plates in the 37 ºC incubator for the next 72 hours. The rest of filtered day 7 supernatants were then stored at -20 ºC until subjecting them to RNA isolation.

72 hours after Ghost cell infection, the Ghost cells were harvested by lifting them with 500uL 1X PBS +5mM EDTA and then centrifuged at 500xg for 5 min. The cells were washed at this point with 1mL of 1X PBS. Cells were spun again and then resuspended in 130uL of basic sort buffer, then fixed with 130uL 4% PFA and vortexed for 10 seconds. Cells were analyzed for their expression of GFP on BD FACSCelesta using BD FACSDiva (software v8.0.1). eGFP fluorescence of the infected Ghost cells was measured using the FITC channel (530/30). Post-

acquisition analysis was then performed using FlowJo software (software v10.4.2). Infection percentage of Ghost cells infected by viruses from u87 cells were calculated and graphed.

On day 14, the viral supernatants were collected, centrifuged at 500xg x 5min and then filtered through 0.45um filter to remove cellular debris and stored at 4 ºC until using later for further infections. The U87 cells were discarded, and a new plate was seeded 6h prior to infection. 6h prior to infection, 100,000 u87 cells were seeded in a 12 well plate. The cells were prepared for infection as mentioned above using 0.5uL of 1000X Polybrene, followed by a 10-minute incubation at 37 ºC. After the incubation, 500ul of the filtered viral supernatants of the LPVs were added to the respective well and mixed by shaking biaxially. The plates were then spinoculated for 1h at 900xg at 25 ºC. The plates are returned to the 37 ºC incubator for 24h. Media containing viruses was aspirated from the cells and refreshed with 1.5mL of complete DMEM ± 5ug/mL PMA, 0.5uM Ionomycin. If PMA/I was added, 1.5mL of complete DMEM was refreshed 24h post-activation. For another set, the virus containing media was aspirated and replaced with 1.5mL of complete DMEM+ puromycin and G418. The remaining supernatant from day 14 was used to check secondary infection in the ghost reporter cells as mentioned above.

This process of refreshing U87 cells was repeated every 14 days up to day 42. And every 7 days, the supernatants were collected to test for secondary infection in Ghost reporter cells. Finally, all the GFP percentages were plotted together to generate a graph showing infection of each LPV in all the timepoints.

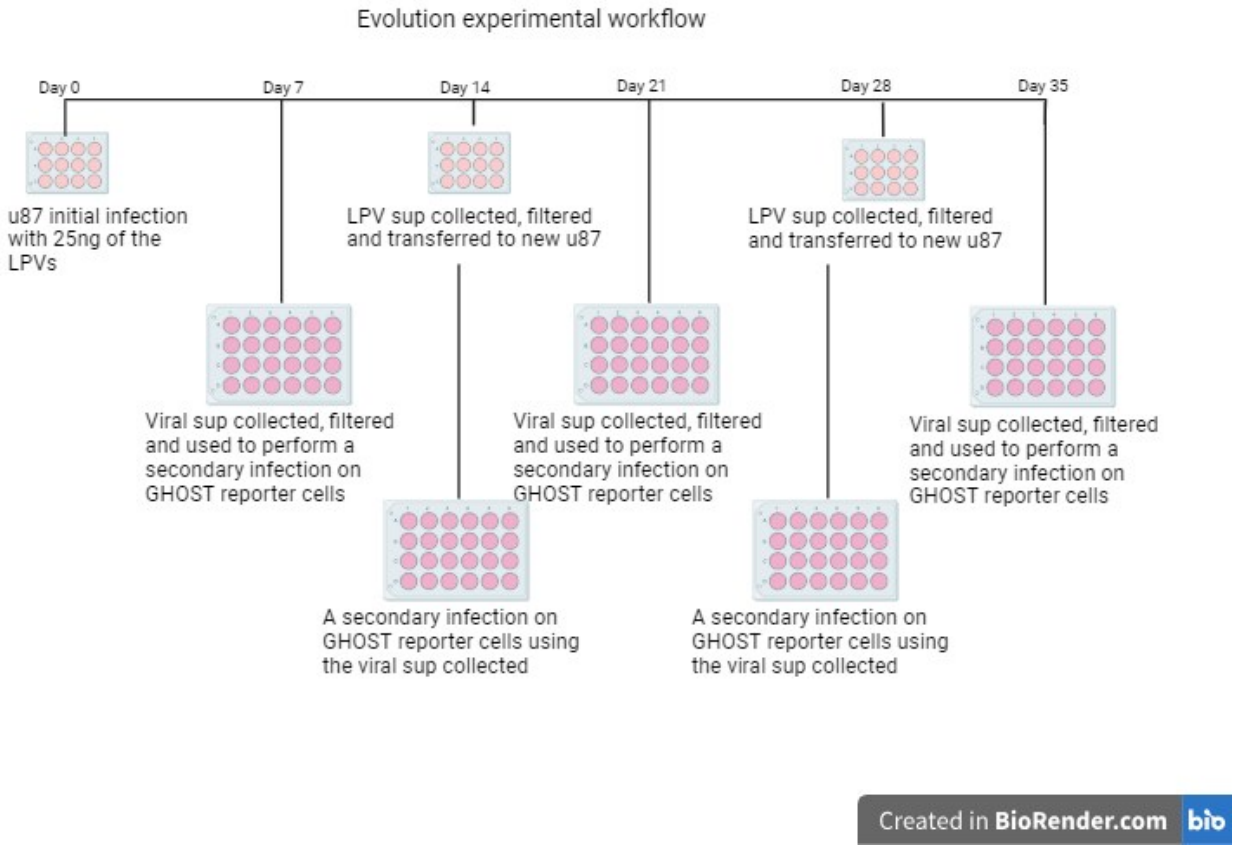


Figure 3.1 Evolution experimental workflow
Protocol for the evolution is pictorially represented. This figure was made using Biorender.

4 RESULTS

4.1 A3 mutations lead to decreased promoter activity indicated by eGFP expression

In order to investigate the activity of the HIV 3' LTR as a promoter, the 3'LTR of A3G and A3F mutated clones were cloned upstream of the eGFP reporter gene. These newly constructed plasmids aided in analyzing the promoter activity of the 3' LTR through functional transfection assays. **Figures 4.1** and **4.2** display the exact location of the A3F and A3G induced mutations. This shows the mutations in clone aligning with their respective TFBS, TAR and other important sites of the LTR for transcription. We evaluated the transcriptional activity of the 3' LTR of a subset of these 57 clones mentioned using two approaches as shown in the **figure 4.3 A-D**. Firstly, we transfected each clone and added PMA/I to half of the samples and DMSO to the other half. The results obtained in **figure 4.3 A** and **4.3 B** provided us with a range of reporter gene expression levels of the various LTR mutants in response to PMA/I. PMA/I, being a protein kinase C agonist, activates the NF-κB and Ap-1 transcription factor pathways. The degree of activation for each clone depends on its mutation profile. Upon examining the percentage of eGFP positive cells, we did not observe a distinct differential pattern between A3G and A3F mutated clones. The zone stratified as "basal" represents the half maximal value of non-induced transcription, while the upper portion of the graph indicates high, and the lower portion denotes low. On the other hand, when analyzing the Mean Fluorescence Intensity (MFI), which represents expression levels as a proxy for transcription efficiency, only 1 clone showed higher MFI compared to the PMA/I induced unmutated wildtype control.

Similarly, a second functional transfection assay was performed using these LTR plasmids with and without the viral protein Tat. The results shown in **figure 4.3 C** and **D** provided us with the range of expression levels of the various LTRs in response to Tat. As Tat transactivation is

essential for efficient viral transcription, the addition of Tat seemed to induce almost all A3F clones and only some A3G clones. Most of the A3G clones remained in the basal zone. The zone stratified as “basal” represents the half maximal value of non-induced, while the upper portion of the graph indicates high, and the lower portion denotes low. The extent of activation for each clone depends on its mutation profile. Similar to the PMA/I activation results, we did not observe a distinct pattern between A3G and A3F in terms of cells expressing eGFP. However, the MFI showed an important difference between A3G and A3F clones. The A3G clones appeared to be poorly induced, with most of them not expressing GFP as intensely as even the non-induced wildtype control. The MFI graphs represent expression levels, as an indirect measurement of transcription efficiency.

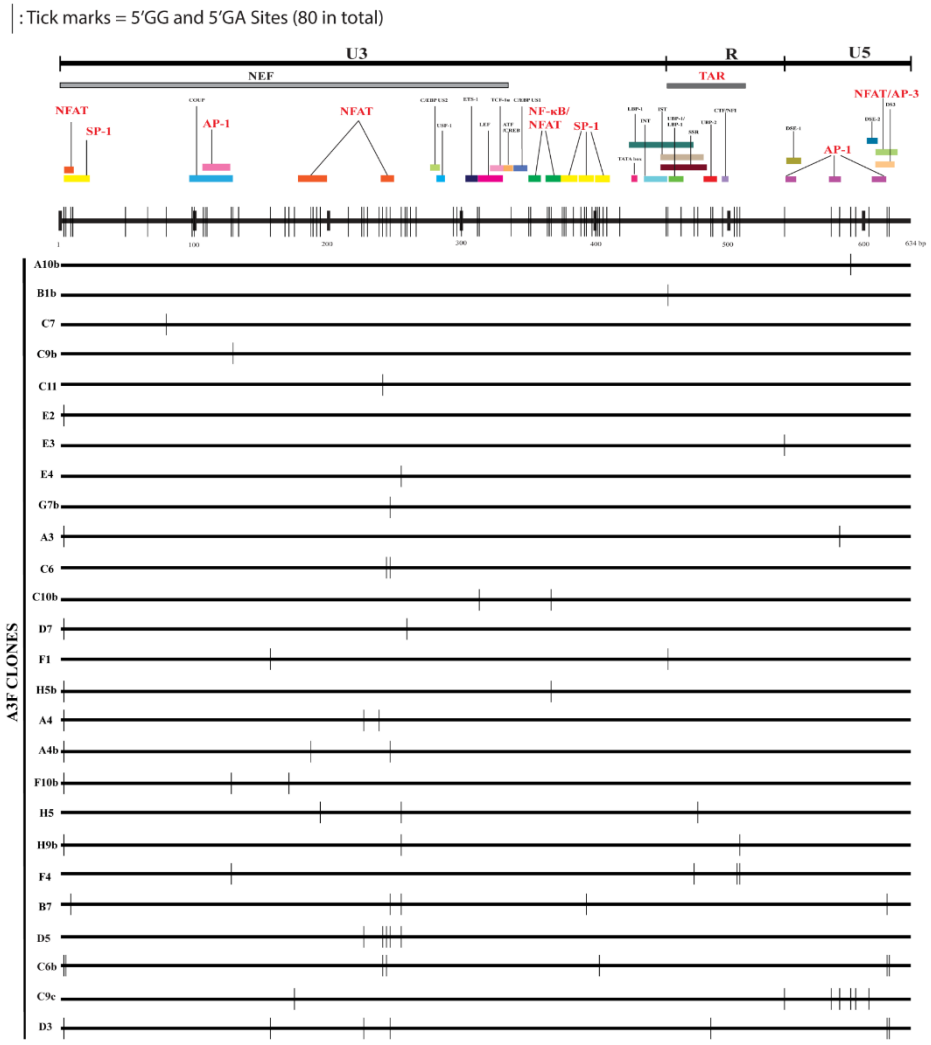


Figure 4.1 Location of the mutations in each clone mutated by A3F.

The tick marks denote 5'GG and 5'GA. This also shows how the mutations align with Transcription factor binding sites.

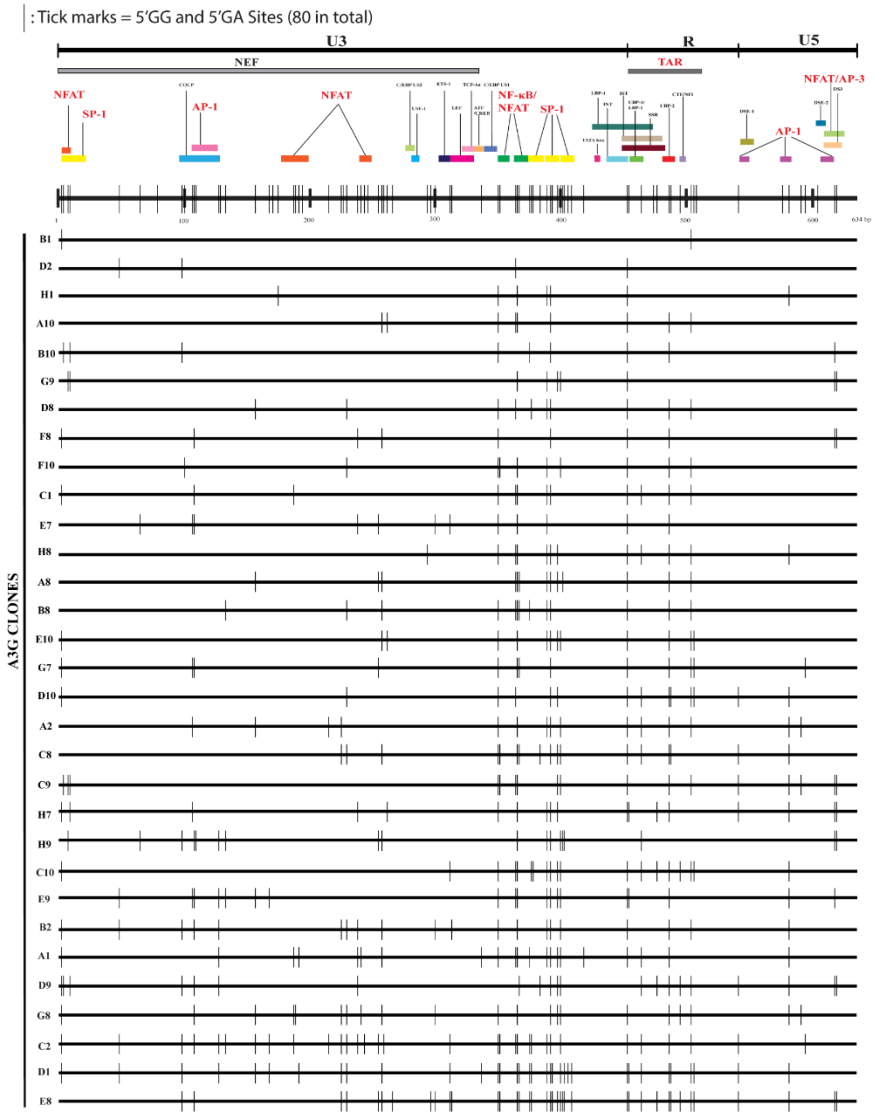


Figure 4.2 Location of the mutations in each clone mutated by A3G

The tick marks denote 5'GG and 5'GA. This also shows how the mutations aligning with Transcription factor binding sites.

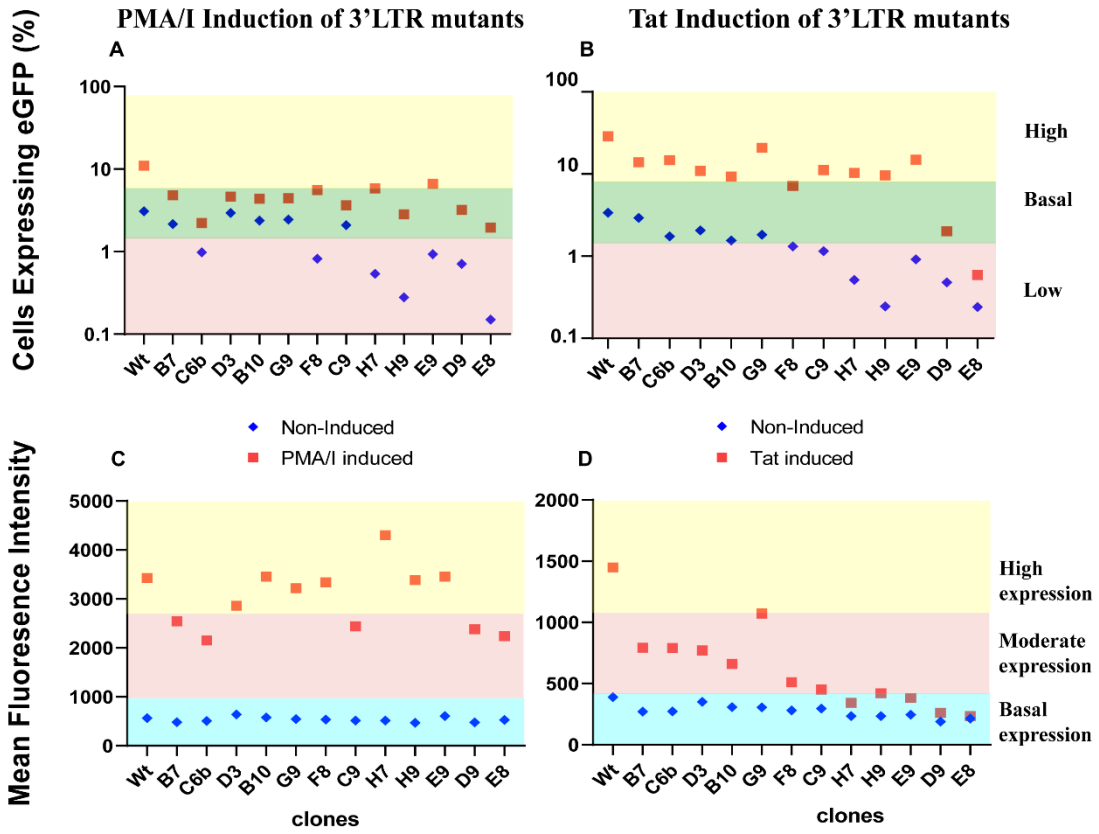


Figure 4.3 PMA/I and Tat induction of 3'LTR mutants

3'LTR mutants induced by by two approaches to measure reporter gene expression by the mutated LTRs as a surrogate for transcriptional activity. Panels A and B represent eGFP expressing cells after PMA/I treatment and their mean fluorescent intensity respectively. Panels C and D represent eGFP expressing cells after Tat transactivation and their mean fluorescent intensity respectively. The zone basal of the panels A and C are half maximal points of the uninduced wildtype. The zones above and below the basal zone are labelled high and low respectively. In the panels B and D, basal expression encompasses the MFI of the uninduced clones and the zones above are stratified as high expression and moderate expression based on the MFI of the induced wildtype.

4.2 NF-κB, NFAT and Sp-1 binding sites are predicted to be affected due to A3 mutations

Our reporter gene expression data supports the idea that the transcription of the mutated LTR is compromised. The goal of this experiment was to ascertain if the compromised transcription is a result of transcription factors being unable to bind efficiently. Here, we used an *in-silico* model to assess the binding affinity of most common transcription factors to the mutated 3'LTR sequences. We evaluated the predicted binding affinity of NF-κB, NFAT, Ap-1, Sp-1, COUP, ETS1 by using an online software called Fabian variant [136]. The result in **figure 4.4** predicts that in most of the mutants, NF-κB, NFAT and Sp-1 binding sites were affected. When comparing across the mutants, we found that the clones E8 and D9 have numerous mutations in several transcription factor binding sites. This could explain why these clones were unable to express the eGFP reporter upon PMA/I induction given that PMA/I activates NF-κB, NFAT and Sp-1 pathways.

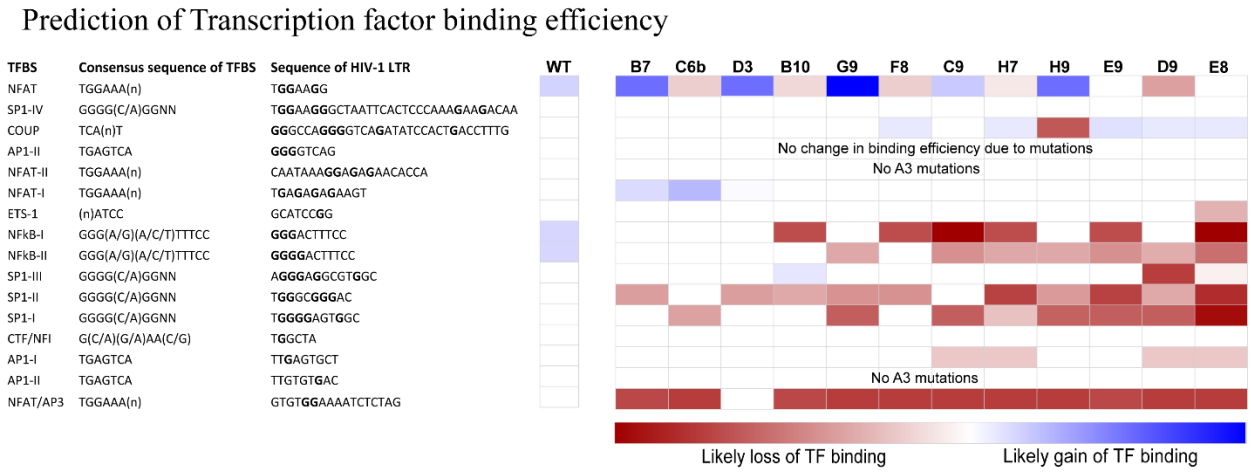


Figure 4.4 Prediction of TFB efficiency

Shows prediction of transcription factor binding efficiency of the wildtype sequence compared to the mutated promoters. This **figure** shows the respective TFBS, their consensus sequence, the sequence on HIV LTR that they bind to and finally the prediction in the form of a heat map. The blue shades indicate gain of TF binding and red shades denote loss of TF binding when compared to the TF binding to wildtype LTR. The white boxes denote that there is no change in TF binding between wildtype and the mutant. The putative mutation sites are written in bold letters under sequence of HIV LTR.

4.3 Non-replicative LPVs show a range of induction to PMA/I

The previous experiments demonstrated that LTR mutations of interest are compromising reporter expression and that this could be due to the inability of transcription factors to bind their respective sites. Next, we wanted to evaluate the effects of the G-to-A mutations on viral transcription, viral infectivity and its response to the latency reversing compound PMA/I, we performed infections using the non-replicative pNL4-3 Δenv, Δvif LPVs. Infections were performed in two different cell lines to demonstrate that the results are not cell specific.

pNL4-3 Δenv, Δvif and the LPVs were produced by transfection and were pseudotyped with VSV-G to enable entry of the virus as these viruses lack env. **Figure 4.5** shows the various levels of eGFP expressed from the viral expression plasmids following transfection. The mutants showed substantially reduced eGFP expression compared to the wildtype. We can see that C6b and D3 are showing better eGFP expression than the wildtype, and LPVs E8, D9, E9, H7 and H9 have very low eGFP expression.

Next, we carried out infections using an ELISA- normalized amount of 345ng of p24 of each LPV and pNL4 wildtype viruses produced in the previous step, in both HEK 293T and Jurkat cell lines. The infections were allowed to proceed for 24 hours before activating the infected cell populations for an additional 24 hours using PMA/I or DMSO as a control. The activation levels were measured using flow cytometry by assessing eGFP expression levels.

Most of the LPVs demonstrated lower baseline levels of eGFP expression following infection when compared to the pNL4 wild-type positive control infecting HEK 293T as well as Jurkat cell lines (**Figure 4.6 A** and **B**). After 24 hours of activation with PMA/I, a subset of LPVs exhibited varying levels of eGFP expression which signifies the infectivity of the respective LPVs. Unlike HEK 293T cells, all clones infected in jurkat cells showed significant increase in infectivity after

PMA/I addition (**Figure 4.6 B**). E8 was the only uninduced clone in HEK 293T cells (**Figure 4.6 A**). Moreover, LPVs with more promoter mutations like E8 and D9 were unable to show inductions as high as LPVs with fewer mutations including B7, C6b and D3. However, some LPVs with higher number of mutations like H9 and E9, provided higher induction in response to PMA/I in Jurkat cells, suggesting that the mutation profile of each LPV determines the ability of viral promoter to respond to PMA/I.

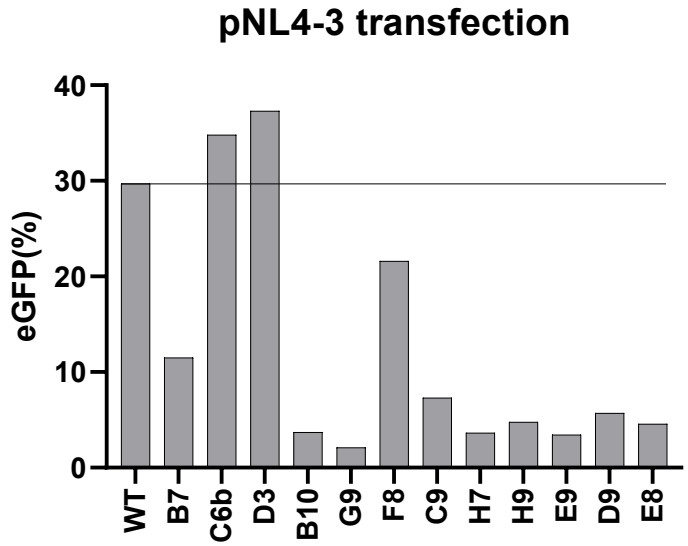
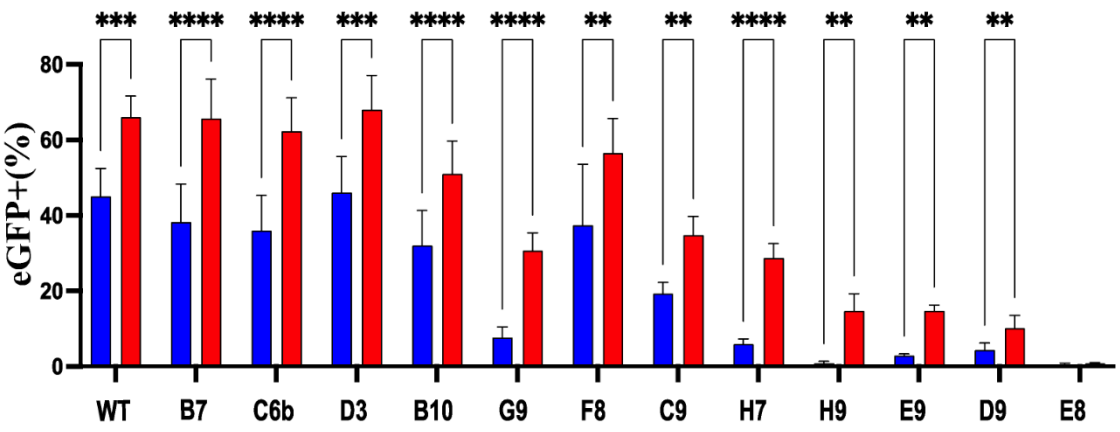


Figure 4.5- Non-replicative pNL4-3 transfection

The LPV plasmids were co-transfected with VSV-G plasmid for the purpose of producing viral particles to be used in in-vitro infection of HEK 293T cells and Jurkat cells.

A- 293T cells statistical analysis



B- Jurkat cells statistical analysis

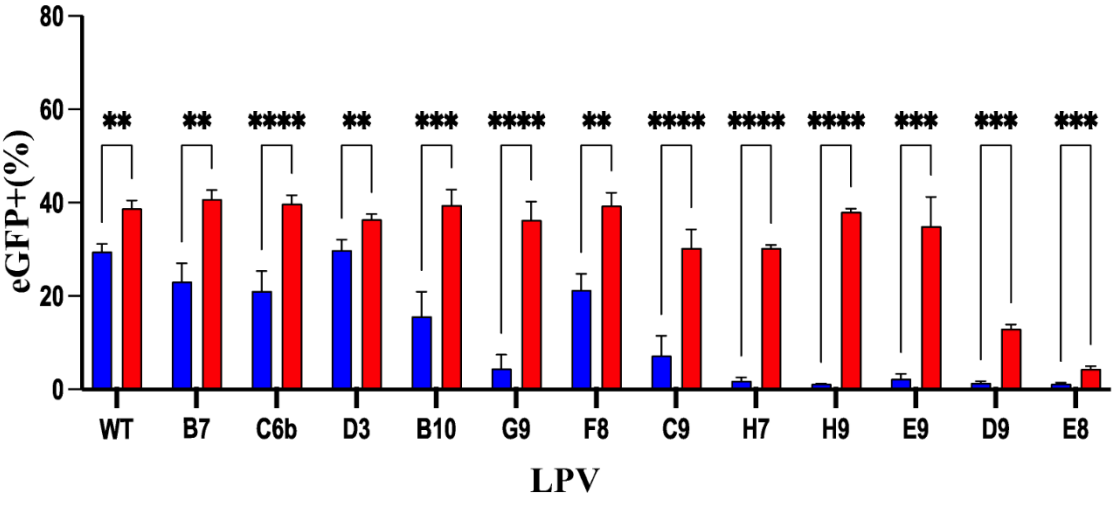


Figure 4.6 Infection with pNL4-Δ*env*-Δ*vif* viruses carrying the 3'LTR mutations

In-vitro infection with non-replicative pNL4-3 viruses and PMA/I inductions in HEK 293T and jurkat cells. Panels A and B represent infection (eGFP%) of HEK 293T cells and Jurkat cells. For statistical analyses of HEK 293T and Jurkat cell infections - Multiple paired parametric T test was conducted. P-value > 0.05 (ns), ≤ 0.05 (*), ≤ 0.01 (**), ≤ 0.001 (***), ≤ 0.0001 (****).

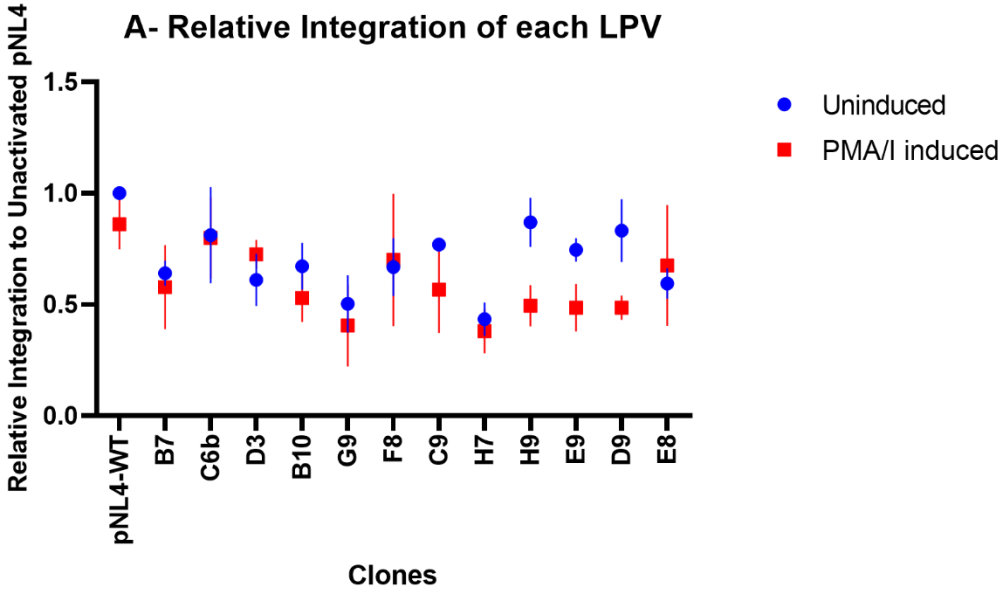
4.4 A3 mutations affect a few of Non-replicative LPVs Integration

To investigate whether the variations in reporter expression are due to different viral integration levels, a semi-nested Alu-qPCR was performed. As described previously, Jurkat cells were infected with the LPVs or pNL4 wild-type and then treated with or without 10ng/mL PMA + 1uM Ionomycin. 24 hours of post-induction, one-third of cell population infected with each LPV was analyzed with the help of flow cytometry to measure eGFP percentages, while the remaining two-thirds of cells had their genomic DNA (gDNA) isolated.

The two-step Alu-qPCR process involved a preliminary Alu-LTR PCR performed on the normalized samples, followed by an internal qPCR on the stage 1 amplicons, so the resulting amplification of integrated virus only. Relative integration levels of the LPVs were calculated by using Actin amplification as initial DNA loading control. LTR amplification data was then compared between multiple trials with an unactivated pNL4 wild-type sample that was constant across all the trials. Finally, the mean integration of the pNL4 wildtype uninduced condition was set to 1.0, and the rest were normalized to this value.

The results showed that there was no significant difference between the integration levels of most of the LPVs when compared to the wildtype virus (**Figure 4.7**). The (**Figure 4.7 B**) displays the statistical analysis carried out to compare integration between each LPV and the wildtype virus. The result indicates that there was no significant difference except for clones E8, H7 and G9 when compared to the integration of wildtype virus. Presence of latency reversing agents upregulate transcription of viral RNA within cells however the drugs do not change viral integration levels (**Figure 4.7 A**). Therefore, neither the LPV mutations nor the PMA/I treatment post-infection seemed to have any impact on the integration of most of latency-prone viruses. As a result, the

differences observed in the LPV infection and PMA/I induction must be occurring due to variations in the ability of these viruses to transcribe and not due to viral integration.



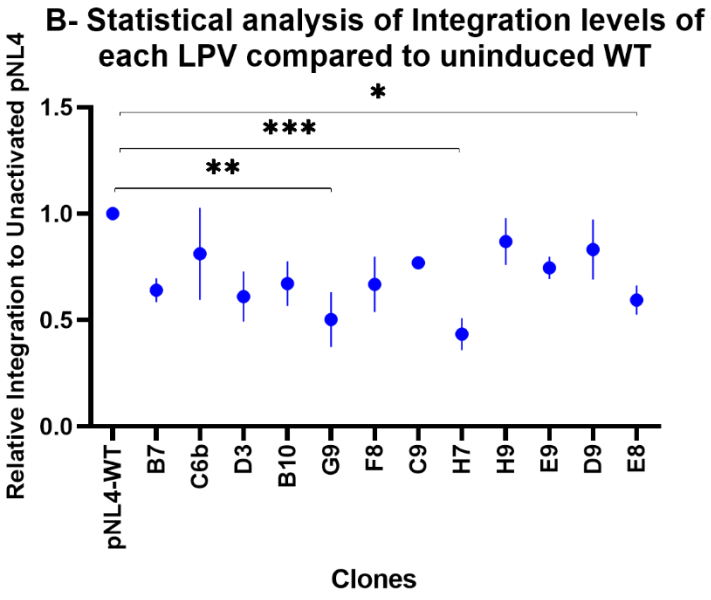


Figure 4.7: **Non-replicative LPVs integration**

Following In-vitro infection and PMA/I inductions of jurkat cells, we measured the relative integration levels of the LPVs compared to wildtype virus. Panels A represents relative integration in PMA/I induced and uninduced LPVs and panel B represents the statistical significance of integration levels of each LPV when compared to the wild type virus. Two-way anova was performed to compare uninduced WT vs LPVs and induced WT vs LPVs. P-value > 0.05 (ns), ≤ 0.05 (*), ≤ 0.01 (**), ≤ 0.001 (***).

4.5 *Ex-vivo* CD4+ T cell infection shows differential infection and induction to PMA/I based on the mutation profile

So far, we have proved that A3 mutations are leading to decreased promoter activity of the LTR and ability of the virus to infect cells. We have also shown that integration levels of most of the LPVs did not have a significant difference compared to the wildtype virus and that the loss of infection is not due to faulty integration but because of the mutations present in the viral LTR. The next part of this objective is to show how these LPVs infect and respond to LRA-PMA/I in CD4+T cells.

Fully replicative CXCR4 tropic LPVs were produced in HEK 293T cells. These viruses were used to infect ex-vivo CD4+T cells isolated from human PBMC. The first step involved in LPV production, where we co-transfected LPV plasmids and VSV-G plasmids in HEK 293T cells. The **figure 4.8** shows the differential eGFP expression of each clone, the effect of mutations is evident right from this step. It shows that the clones C6b, D3 transfect similar to wildtype, other clones B7, G9, F8, C9, H7 and H9 transfect less efficiently as determined by eGFP expression in the transfected cells (**Figure 4.8**) and hence to have 2000ng equivalent of p24 we needed to concentrate the virus 4-20x times. The rest of the clones E8, E9, D9 have many promoter mutations that ultimately produced very low viral titre even when concentrated 30x times. Surprisingly, the clone B10, which does not have as many mutations as E8 or E9, but we still couldn't produce enough of this LPV to proceed with infection (**Figure 4.8**).

CD4+ T cells were isolated from human PBMCs using an isolation kit by negative selection and the isolated cells were subjected to a 72h activation with PHA and Il-2. Then, activated T cells were infected with the fully replicative LPVs equivalent to a high p24 mass (2000ng) in-order to discover the maximum potential of the LPVs. They were then treated with PMA/I. Post

experiment, the cells were collected and externally stained for CD4, CD3 and CD45 which are the basic markers of a CD4+ T cell. Following that, a few infected cells were internally stained for p24 after permeabilization of the cells with a p24 antibody tagged with the fluorophore APC, and a few were stained without permeabilization to check for internal and external p24 expressions respectively. The cells were later analyzed by flow cytometry to determine the infection and induction to PMA/I. **Figure 4.9** shows the gating strategy and back-gating explaining how the cells were selected and analyzed.

Similar to the results obtained with in-vitro experiment, we observed differential inductions and infections with different clones while infecting CD4+T cells with the replicative LPVs (**Figure 4.10**). Few LPVs demonstrated observable lower basal infections as measured by eGFP expression compared to the pNL4 wild-type control like C9, H7 and H9. After treatment with PMA/I, some of the clones showed better infection as measured by eGFP% and p24% separately, and some clones activated similar to wild-type virus and some others did not induce. **Figure 4.10 A** shows the infectivity of the LPVs in terms of eGFP%. H9 is a clone that was showing the best induction in the in-vitro experiment, it is assuring to note that H9 has the best induction in this experiment as well. This tells us that the mutations present in H9 promoter enables it to be induced better than the rest of the LPVs. Apart from clone H9, B7 and C6b show significantly prominent induction when treated with PMA/I, whereas clones G9, F8, D3 show mild induction that is not significantly better than the uninduced. Clones C9 and H7 are the least inducible LPVs showing induction by just 1% -3% (**Figure 4.10 A**). **Figure 4.10 B** shows the infectivity in terms of viral protein p24 expression, as it is one of the best-known diagnostic parameters for HIV infection. It was interesting to observe that PMA/I treatment had a significant impact on p24 expression of only the wildtype virus, non-significant but observable increase of about 10% in P24 expression of LPVs

C6b and H9 (**Figure 4.10 B**). A few LPVs including D3 and F8 showed an induction of about 5% in p24 expression. And a few other LPVs including G9 and H7 showed no significant difference in p24 expression between uninduced, and PMA/I induced cells. Lastly, we compared the p24 expression in infected cells by internal and external staining of p24 (**Figure 4.10 C**).

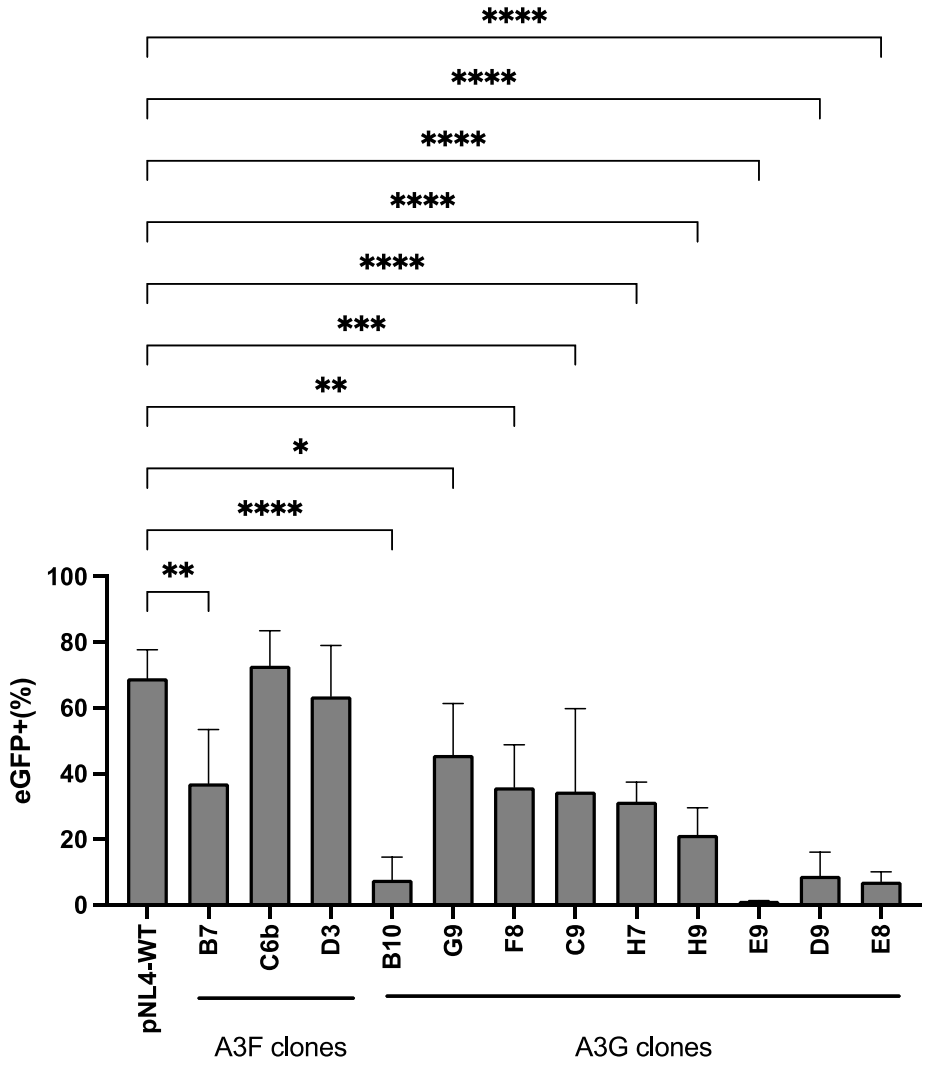


Figure 4.8- Expression of eGFP cells transfected with replicative CXCR4 pNL4-3 LPVs.

The LPV plasmids were co-transfected with VSV-G plasmid for the purpose of producing viral particles to be used in ex-vivo infection of CD4+ T cells. Multiple paired student T test was conducted. P-value ≤ 0.05 (*), ≤ 0.01 (**), ≤ 0.001 (***), ≤ 0.0001 (****).

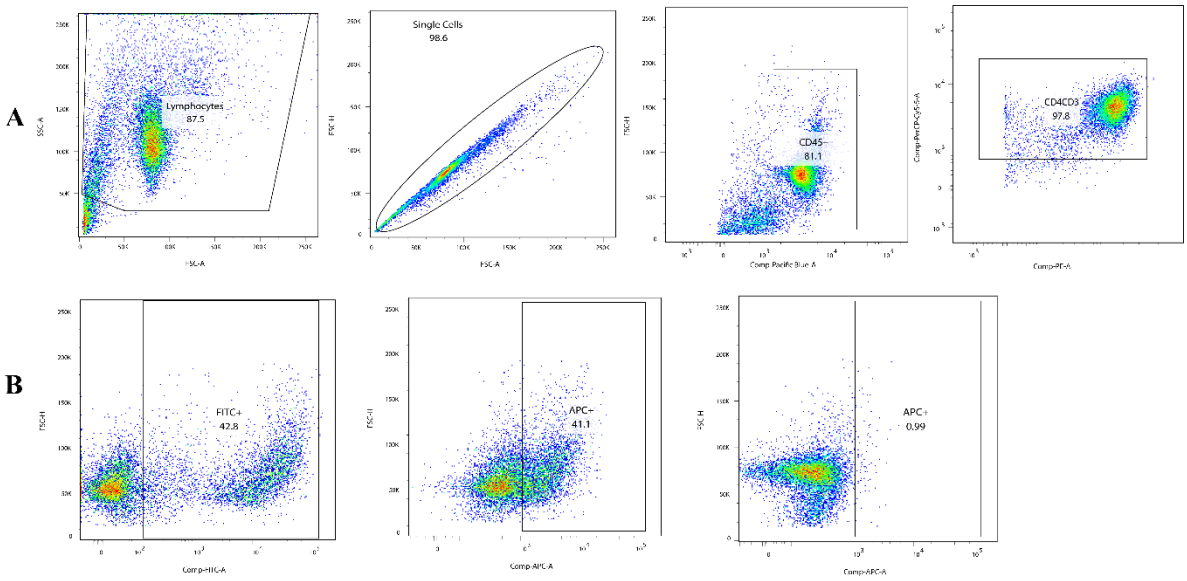


Figure 4.9 Gating strategy used to analyse infected CD4+ T cells

A. Shows how the CD4+ T cells that were infected with wildtype virus were selected with an overview of the gating performed. Each graph determines the percentage of cells expressing the respective marker.
B. Shows the percentage of cells expressing eGFP, internal p24 and external p24.

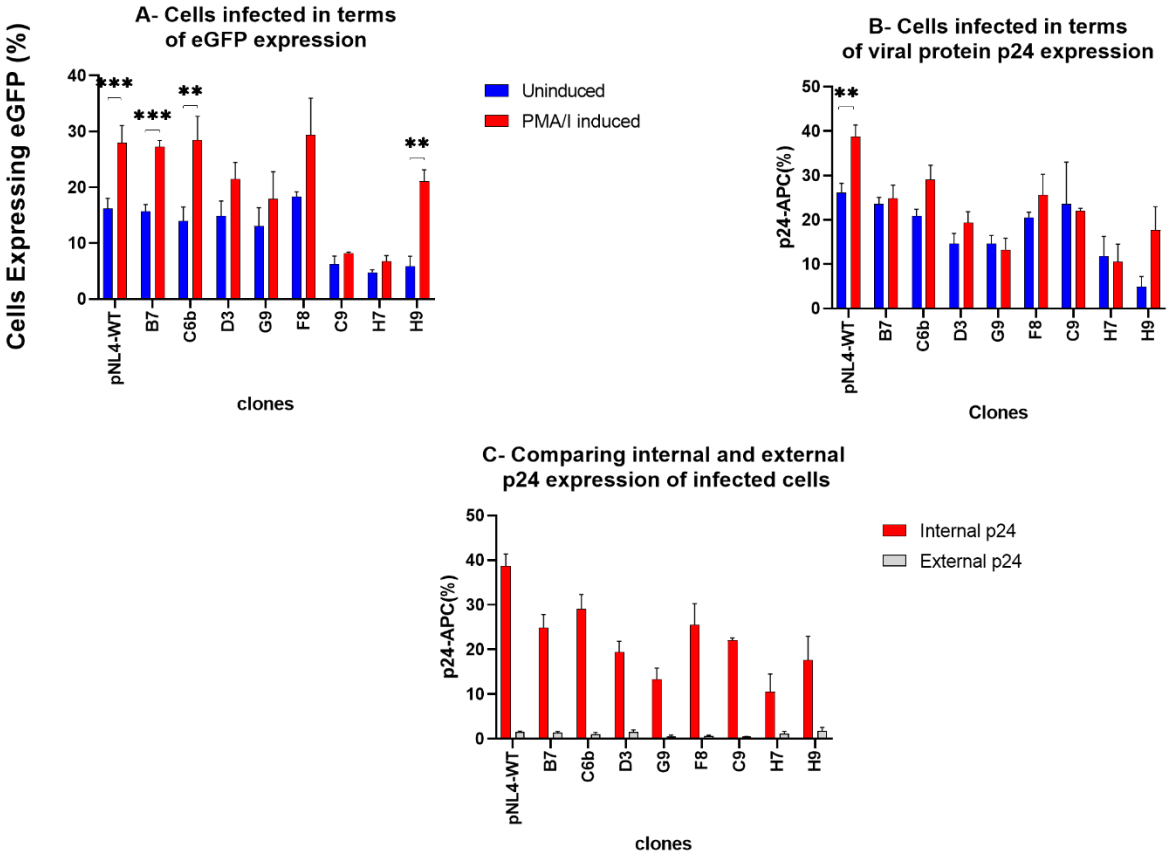


Figure 4.10: **Ex-vivo infection and PMA/I inductions of human CD4+ T cells**
Panels A and B represent infection in terms of cells expressing eGFP and viral p24 respectively. Panel C shows the difference between cells that internally express viral protein p24 vs external detection possibly due to cytotoxic activity. Multiple paired parametric student T test was conducted. P-value > 0.05 (ns), ≤ 0.05 (*), ≤ 0.01 (**), ≤ 0.001 (***).

4.6 LPVs C6b and B7 reverted in response to PMA/I

The final objective is to determine if the LPVs can revert to a more transmissible state following with multiple rounds of infections and inductions. To attain this, U87 CD4+ CXCR4+ cells were infected with 25ng of p24 of each replicative LPV supernatant, that were VSV-G pseudotyped to enable access into the U87 cells, which are acting as viral reservoirs producing latent viruses. The initial viral supernatant was aspirated, and the cells were then induced for 24 hours with PMA/I. Every 7 days, supernatants of these infected cells were harvested, and the percentage of replication-competent viruses released in these supernatants was measured by performing an infection on Ghost reporter cells. Ghost cells are eGFP reporter cells carrying HIV-2 LTR which is capable of producing eGFP when viral protein Tat is produced.

In **Figure 4.11**, we are showing results from 4 separate trials of this experiment. In the first 14 days of all trials, B7, C6b LPVs of almost all trials were able to produce detectable levels of replication-competent viruses, as measured by expression of eGFP by the Ghost reporter cells. At day 21 some of these clones shut down. In subsequent experiments, LPVs D3, F8 and G9 also exhibited detectable, infectious viral rebound between day 14 and day 21. In certain instances, the eGFP levels of a few LPVs (especially C6b) reached that of the wild-type control, and in some others, the virus rebounded very briefly (examples are D3, B7). There were cases where, some LPVs disappeared around day 21.

The LPVs B7, D3, and C6B have the least number of LTR mutations, and this could be a reason for these clones to revert. This hints us that if given more time some of the other LPVs could bounce back and thus our experiment demonstrates that viruses exhibiting A3 introduced mutations are replication-competent and capable of reseeding an infection given enough time.

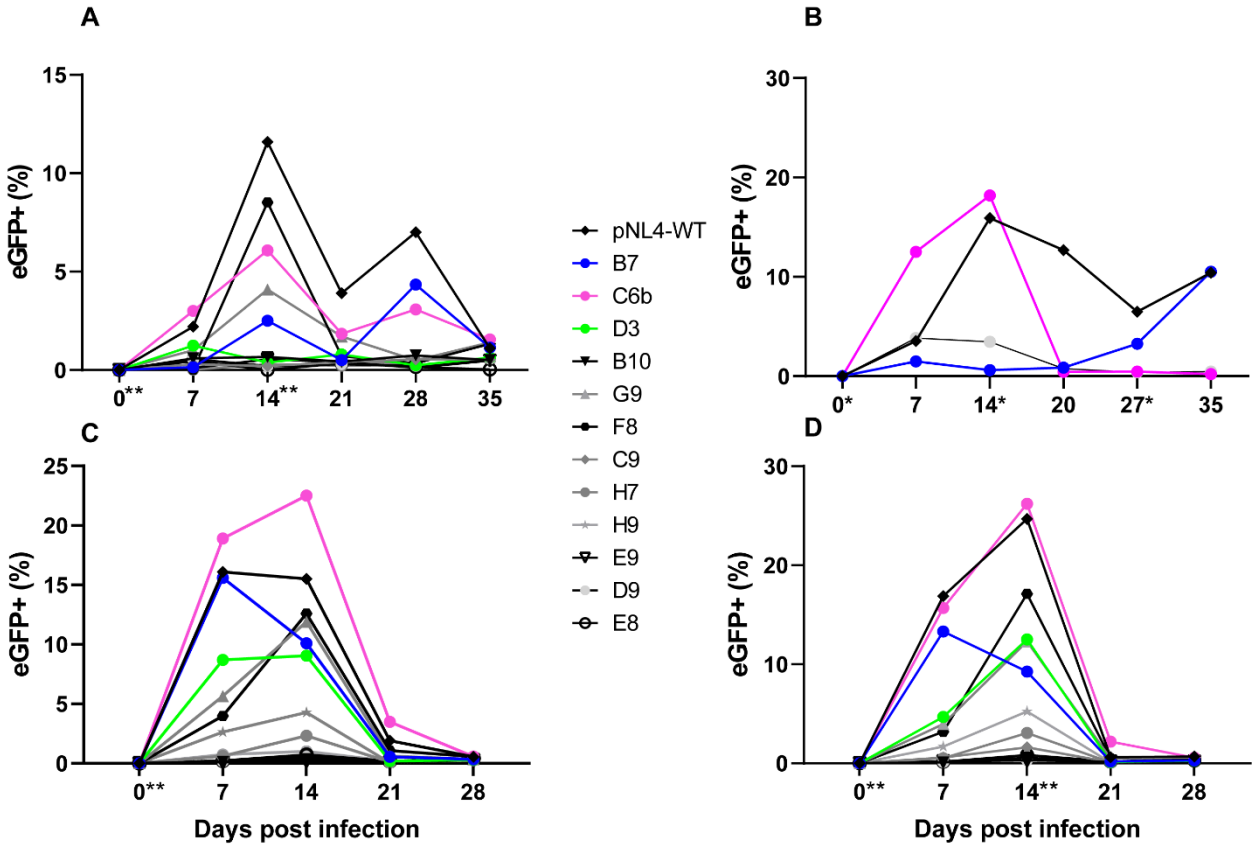


Figure 4.11- Reversion of the LPVs from a state of transcriptional silence
Viruses produced from a grow-out experiment were used to infect ghost reporter cells. Viral output is measured as eGFP % by flow cytometry. PMA/I was added at different timepoints in each experiment, "**"represents time of PMA/I addition.

5 DISCUSSION

HIV is known to persist in the cells it infects, as latent reservoirs. While continuous efforts have been made to treat the disease, HIV infection is still incurable and there is currently no vaccine against it [20, 124, 137]. Antiretroviral therapy regimes undertaken by PLWH help in reducing the spread of HIV infection *in-vivo*, although the persistence of already existing latent reservoir requires treatments throughout life[138]. The ability of the virus to persist in latent reservoirs has protected them from the host immune system, and additionally, antiretroviral therapies are unable to reach the reservoir [29]. Latent reservoirs are established during the early phases of HIV infection, primarily infecting the effector CD4+ T cells while they transition to memory T cell phenotype [4, 32, 139]. Some other cells that are shown to harbour latent reservoirs are macrophages, monocytes, and dendritic cells in various anatomical locations[32, 139].

The establishment of latent reservoirs in an infected individual occurs due to various epigenetic and non-epigenetic mechanisms as discussed in the introduction [29, 66, 70, 130, 140]. In addition to the commonly identified mechanisms mentioned above, we are interested in a new mechanism called viral molecular latency. Here, we propose that the mutations in the viral genome cause a state of latency. Ho and co-workers [27] studied various latent reservoirs and discovered that most of the latent viruses have defective genomes and concluded that the defective genome does not result in a productive infection. However, that is most likely when there are large numbers of mutations while small numbers of mutations could result in non-lethal effects aiding the persistence of the virus. Latent HIV reservoirs are characterized by integrated and replication-competent viruses but temporarily being transcriptionally inactive [27, 29, 66, 83].

APOBEC3 proteins are host produced antiretroviral proteins that have the capability to introduce G-to-A transition mutations on the HIV genome leading to generation of premature stop codons

[82, 107]. However, our lab is interested in mutations introduced in the HIV LTR which is also the viral promoter. We know that HIV produced viral protein Vif, is capable of ubiquitinating and degrading the A3 proteins [105] so, as the disease progresses A3 levels packaged in the egressing virus decreases. We believe that, the low levels of packaged A3 creates non-lethal mutations in the viral genome leading to production of viruses displaying a latency-like phenotype. So, our hypothesis states that, G-to-A mutations introduced in the HIV LTR by RT error and APOBEC3G/F mutations result in a latency-like phenotype.

To address this hypothesis, a library of clones having mutations in their 3'LTR was generated (by Ms. Cindy Lam, one of the predecessors in our lab) by co-transfecting HIV plasmids with varying amounts of A3G/F (refer appendix 8.1). Thus, a library of 57 clones of HIV having about 1 to 38 mutations in their 3' LTR region was generated. Out of the 102 GG/GA putative A3 mutation sites, 80 sites exhibited mutations. **Figures 4.1 and 4.2** show the exact locations of mutations in all the 57 clones. Twelve clones were shortlisted from the set, based on how they responded to Tat stimulation (3 from A3F mutants and 9 from A3G mutants, hence named as A3 mutants). These mutations were previously crosschecked with HIV patient sequences (refer appendix 8.3). Several of the mutations in the TFBS we identified were also found in HIV patient sequences. The clones that showed subtle induction when compared to the wildtype were selected, as we expect these clones to show a latency-like phenotype. The 3'LTR which bore the mutations of interest were cloned into 3 types of plasmid backbones by Ms. Cindy lam and Mr. Matthew Greig, predecessors in our lab. The mutated 3' LTRs were cloned into pEGFP-C3 backbone (to assess the promoter activity of the LTR) and were also cloned into replicative and non-replicative viral plasmid backbones (to measure viral infection and induction to PMA/I treatments).

5.1 Measurement of Promoter activity

The initial question we asked was, if the A3 induced mutations in the 3' LTR region, affect their promoter activity. From the functional transfection assays performed using the 3'LTR mutants, we discovered that the uninduced LTR transcription was hampered (**Figure 4.3 A and C**) which was indirectly shown by the basal eGFP expression. The two routes of induction, PMA/I and tat treatment that we performed on the mutants showed that, the mutants induced differently to different treatments (**Figure 4.3**). Some of the clones responded better to PMA/I rather than tat when comparing their fluorescent intensities, as it indirectly signifies the transcription efficiency (**Figure 4.3 B and D**). But the number of cells expressing eGFP after tat transactivation does not appear to be compromised (**Figure 4.3 C**). The reason for this could be associated with the mutation profile of the clones in question. The A3G mutants, E8 and D9 have the maximum number of mutations. In particular, they have multiple mutations in the TAR region on top of several mutations in the transcription factor binding sites. Hence, these mutants were not transactivated by tat treatments. The sequence of the TAR region in the clones E8 and D9 have nine and five mutations respectively, so it is possible that the viral protein Tat was unable to bind and transactivate. Interestingly, clones H9, H7 have one and four mutations respectively in the TAR region, and both the clones are not induced by treatment with Tat. On the other hand, the clone D3, also has one mutation in the TAR region, but Tat was able to induce eGFP expression. The reason for this could be mutations in rest of the LTR. The clone D3 is one of the A3F mutants having relatively low mutations, hence it was able to salvage and respond to treatment with Tat. The other clones track with their mutation profile and the treatments. Each point on the graph (**Figure 4.3**) is an average of two trials, hence we have not analysed its statistical significance.

The drawback of this experiment is that we are measuring transcription in terms of eGFP expression which might not relate to the LTR transcription always. A solution for this would be to perform a RT-qPCR [141] to determine the transcription of the mutated LTR.

The key points worth mentioning from the first experiment are, our mutations are modulating the LTR eGFP expression that can be in-directly linked to its transcription capability with and without the PMA/I and tat treatments. Based on observations from both PMA/I and tat treatments and the mutation profile, the clones B7, C6b, D3, B10, F8, G9, H7 and H9 have the highest possibility of displaying a latency-like phenotype in a viral environment.

5.2 Analysis of Transcription factor binding to the mutated promoter

The previous experiment proved that the mutations led to decreased eGFP expression hence low transcription is expected to occur. We now asked if the anticipated low viral transcription is due to inability of transcription factors to bind the mutated TFBS. To address this question, we performed a pilot study with the help of an online tool called Fabian variant [136]. This experiment (Figure 4.4) predicted that, the NF-κB, NFAT and Sp-1 binding sites were compromised due to the mutations of interest. Since there are multiple binding sites for each transcription factor, some of them being compromised will only affect a little of its overall transcription ability. The clones B7, C6b and D3 have a predicted gain in the binding of NFAT at a few sites but they also have loss in transcription factor binding in binding sites of other transcription factors which could affect the overall transcription of the LTR. The transcription expected in each clone is based on how important each binding site is for overall transcription to occur. Research has shown that NF-κB, NFAT and Sp-1 are vital for HIV transcription [3, 14, 143]. Clones E8 and D9 were predicted to have compromised TF binding in all the mentioned sites and hence, we expect these clones to be totally non-infective. There are some clones having one of the two NF-κB binding sites

compromised and we expect these clones to display a slightly better phenotype than the E8 and D9 in terms of their capacity to infect. Considering all the results of past and present experiments, we are inclined to predict that the clones B7, C6b, D3 that do not have many TFBS compromised, B10, G9, F8 and H9 that have one of either NF-κB or Sp-1 sites compromised, have the highest probability to not have a massive block in transcription due to TFB. Since this is an *in-silico* experiment, it is just a prediction and this further needs to be evaluated by performing CHIP assays[144] or microarray [142] to verify the loss/ gain of transcription factor binding.

5.3 Measurement of infection and PMA/I induction of the LPVs

Now we know that the mutants B7, C6b, D3, B10, G9, F8 and H9 are of particular interest because of how they responded to the previously discussed treatments and additionally, the TFBS prediction also ended up predicting the same clones. Next, we asked how these mutations affect viral infection when present within a viral environment. To attain this, we first produced viruses using the non-replicative viral plasmids lacking *env* and *vif* genes that were co-transfected with VSV-G to produce the non-replicative viruses that are capable of allowing only one round of infection. **Figure 4.5** shows the differential eGFP expression of transfected 293T cells which resulted due to the mutations. This evidence shows the extent of effect each mutant could have on the LPVs produced as a result of their mutation profile. The viruses were titrated by ELISA and calculated volumes of LPV supernatants were utilised to perform infections in HEK 293T and Jurkat cell lines. HEK 293T cells were first used because of the pseudotyped non-replicative viruses. Whereas, Jurkat cells were chosen as they are a T lymphocyte cell line with high susceptibility to HIV infection because of production of high levels of IL-2 and J-Lat is a most commonly used HIV latency model which is a Jurkat cell line [145, 146]. PMA/I treatment was carried out 24h post-infection for 24h after which the cells were collected to be analysed. Using

flow cytometry, the capability of each LPV to produce HIV specific viral proteins in the presence or absence of PMA/I was measured through expression of eGFP.

As per our initial observation, all LPVs demonstrated comparatively lower percentages of basal eGFP expression to the pNL4 wild type, positive control (**Figure 4.6**). Next, the eGFP expression increased when treated with PMA/I (**Figure 4.6**). The trend observed between the two cell lines was similar. However, the Jurkat cells are more reliable because they are susceptible and permissive to HIV 1 infection and facilitate comparison of infections between several HIV LPVs [146]. The LPVs that were mutated by A3G including C9, G9, H7, H9 and E9 exhibited basal eGFP of around 0% which was induced by PMA/I treatments up to eGFP levels similar to induced wildtype virus (**Figure 4.6 B**). Apart from LPV E8 all the LPVs in the HEK 293T cells resulted in a significant increase in the viral infection in response to PMA/I (**Figure 4.6 A**). Surprisingly, all the LPVs infecting Jurkat cells were able to create a significant increase in infection after PMA/I treatment (**Figure 4.6 B**).

It was interesting to note that LPVs with higher number of mutations usually showed the least percentages of eGFP in infected cells including the clones E8, D9, H7, H9 having their basal infection at 0% (**Figure 4.6 B**), possibly due to the assumption, more the number of mutations in an LPV promoter, greater will be the number of transcription factor binding sites being compromised. Loss of transcription factor binding results in a reduced eGFP expression as predicted in the previous experiments (**Figures 4.3 and 4.4**). Integrating this result with the previous one, we noticed some of the LPVs to show significantly increased response to PMA/I (**Figure 4.6 B**), possibly acting differently due to their mutation profile. We know that NF-κB, NFAT and Sp-1 are very important for viral transcription [14, 73, 112], and the combination of mutations leading to compromised transcription factor binding could interfere with the ability of a

virus to transcribe. The LPV H9 has only one of its NF-κB binding sites compromised (**figure 4.4**), hence was able to respond significantly better to PMA/I treatment (**Figures 4.6 A, 4.6 B**). Similarly, the clones B7, C6b D3 that had a comparatively better basal infection also showing significantly increased response to PMA/I possibly due to having fully functional NF-κB or Sp-1 binding sites (**Figures 4.4, 4.6**). On the contrary, the LPVs H7, C9 and E9 show increased induction even though they have most of their NF-κB, Sp-1 and NFAT binding sites compromised. The reason for this event could address the possible drawback to this experimental set up only allowing one round of infection. Overall, we believe that the LPVs B7, C6b, D3, B10, F8 and H9 are displaying features consistent with being in a latency-like phenotype so far with all our observations.

5.4 Analysis of Integration levels of A3 mutants

The LTRs of HIV have significant role in the integration process of the virus, so the mutations within them might have an effect on their integration efficiency [54, 81, 147]. Even though we have shown that the mutants are giving rise to viruses with compromised infection and ability to respond to PMA/I, literatures show that APOBEC3 proteins are also associated with altering HIV integration site profiles [82], thereby affecting their integration levels. Hence, we checked with this experiment if the mutations introduced by A3s have any role in viral integration meaning, to verify if the differences we observed with the infection and activation of the non-replicative LPVs were due to differences in their transcriptional response or difference in their integration levels.

In order to address our questions, we performed a two-step Alu-qPCR on the cellular genomic DNA (gDNA) extracted from Jurkat cells that were infected with our non-replicative viruses and treated with/without PMA/I. Alu elements are the most abundant and the most repetitive elements in the human genome [148, 149]. All the LPVs in a pcr run were compared to actin levels first and

were normalized to the wild type (WT) pNL4-3 unactivated sample that was set to 1. So, the integration levels of all the LPVs were calculated with respect to WT unactivated control.

Figure 4.7 compares the mean normalized integration levels of all LPVs, and there seemed to be no significant difference between most of the LPVs when compared with wildtype. In **Figure 4.7 B,** the uninduced LPVs H7, E8 and G9 were showing significantly lower integration when compared with uninduced wildtype. Similarly, PMA/I induced LPVs H7 and E8 were integrating significantly lower than the induced wildtype virus (**Figure 4.7 A** statistics not shown). These clones have multiple mutations in the beginning and end of the LTR which could potentially alter their ability to integrate with the host genome. Therefore, the difference in eGFP expression denoting the infection and PMA/I induction percentages among the LPVs when treated with PMA/I were likely due to the ability of the individual promoter in response to PMA/I, and not due to differential LPV integration levels in most of the LPVs except E8 and H7.

5.5 Measurement of infection, replication, and PMA/I induction of the LPVs

For the next phase of this project, we asked what happens if the LPVs are able to undergo more than one round of infection. Firstly, we asked how these mutations affect viral infection when undergoing more than one round of replication. We used the third type of plasmid which if fully replicative and composed of the mutated LTRs in a CXCR4 tropic backbone cloned by a previous student Matthew Greig (Appendix 8.2). To attain this, we first produced viruses using the prior mentioned LPV plasmids. **Figure 4.8** shows the differential eGFP expression resulted due to the mutations. Similar to the non-replicative transfection, this figure also shows the same pattern for various LPVs. Statistical characterization showed that all the LPVs were produced significantly lower than the wildtype virus. The titre of the viruses was very low hence some of the LPVs were concentrated upto 30x times to produce enough viruses for infection. However, the LPVs E8, E9

and D9 were heavily mutated and hence couldn't produce enough of and was not sufficient for infection even after 30x concentration. These heavily mutated clones were expected to be less infectious as shown in previous experiments in this thesis, however it was surprising to see that the LPV B10, that had was able to respond to treatments in the previous experiments turned out to produce very less viral titre (**Figure 4.8**). So, the above-mentioned clones were eliminated from this experimental setup. The plasmids coding for replicative virus were co-transfected with VSV-G to produce the fully replicative CXCR4 tropic viruses and used to infect human CD4+ T cells. The infection and inter p24 characterization were done using flow cytometry and **Figure 4.9** shows the gating strategy used to obtain the graphs presented in this experiment. The very first observation here was that a few of the LPVs exhibited comparatively lower levels of basal eGFP and p24 expression to the wild type, pNL4 positive control (**Figure 4.10 A and 4.10 B**) which is due to mutations residing in transcriptionally important sites thereby affecting viral infection. Secondly, the eGFP expression increased only for a few LPVs when treated with PMA/I (**Figure 4.10 A**). The LPVs B7, C6b, D3 and H9 showed significantly higher eGFP expression compared to their basal infections. However, the p24 expression of the LPVs after PMA/I induction did not show a statistical significance but had a subtle increase in the viral protein expression only in B7, C6b, D3, F8 and H9 (**Figure 4.10 B**). The LPV F8 does not show a significantly higher eGFP after PMA/I treatment but, the difference is pronounced more than the other LPVs in this experiment that were not induced (**Figure 4.10 A**).

It was interesting to note that the LPVs with the lowest number of mutations showed an increased p24 expression in alignment with their eGFP expression following PMA/I treatment (**Figures 4.10 A and B**), examples of this statement are the LPVs B7, C6b and D3. Similar to the *in-vitro* experiment, the LPV H9 has the highest induction, because PMA/I able to compensate the effects

of the mutations during the infection [150, 151]. It was however interesting to note that the LPVs C9, H7 and G9 expressed more p24 when comparing to the eGFP expression in the same cells. The possible reason for this lies in the inability of these LPVs to undergo complete productive transcriptions to express eGFP, but are able to produce viral protein p24 as the open reading frame responsible for producing p24 lies in the early regions of the viral genome [3, 19, 152].

Since these are primary cells, we don't have the luxury of characterizing the latent reservoir and how long this takes for certain types of LPVs to revert. Overall, we believe that the LPVs B7, C6b, D3, F8 and H9 are displaying features consistent with being in a latency-like phenotype after including the last result.

5.6 Determining if the LPVs exhibit reversion

Finally, we asked if the LPVs evolve and revert to an active replicative state if allowed to grow over a few weeks as that is one of the most important features of viruses exhibiting latency[29, 153]. We present results from four different experiments (**Figure 4.11**). The clones that showed signs of reversion in all the trials are B7 and C6b. This is in accordance with predictions and results shown throughout this thesis. In the first 7 days of infection, a few LPVs showed subtle secondary infection. It was interesting to look into LPV C6b as it showed the highest infection in all the 4 trials at day 7. At day 14 the LPVs reached its peak and moving forward they started declining. The same observation is documented in patients as the viral load starts to decline after the phase of acute infection[29, 154, 155]. In addition to LPVs B7, and C6b, D3, F8, and G9 also attained production secondary infection in Ghost cells at day 14. Initial experiments were carried out with and without PMA/I treatments, (data not shown) so as to compare if the clones are able to revert faster with the drug than without. But we realised that adding PMA/I once or twice during the

experiment is to basically push the viruses to start replication and there was no benefit in comparison.

B7 and C6B are A3F-mutated viral clones and contain the least number of LTR mutations out of the 12 LPVs tested here, so it is understandable that these LPVs were able to rebound in infectivity in all the trials. In addition, out of all the LPVs tested these do not have any mutation in the TAR region. The TAR region facilitates Tat binding to initiate and elongate the viral transcription. When Tat does not bind TAR, the efficiency of HIV transcription decreases as RNA pol II fails to move forward [74, 121]. Surprisingly, the clone D3 although having very less mutations in the LTR only showed reversion in one of our trials. If compared with previous results in this thesis, this clone seemed to be always showing higher infection percentages. The reason for this clone not being as consistent with this experiment could be due to one of its mutations present in the TAR region. If the mutation in TAR prevented its reversal, then it would be interesting to check the sequence of the virus that reverted. Studies have discussed that HIV evolves with a combination of mutations in its core promoter, viral enhancer, NF-κB and Sp-1 transcription factor binding sites, as shown in our LPVs [156]. Moreover, mutations in TAR region have been determined to deter viral transactivation and translation [157, 158]. So D3 that rebounded could be due to extra mutations that were introduced in the genome that basically counteracted the mutation in TAR as it is highly unlikely that the viral rebound occurred due to a direct reversion of specific G-to-A transition mutations that were originally introduced. It is not through further APOBEC3 exposure that this reversion could occur as A3 proteins solely create G-to-A transition mutations in the +ssDNA strand, and not A-G. Furthermore, the odds of a perfect reversion of the exact mutation are very low even though it could occur through RT activity. So, it is possible that newer mutations were accumulated outside of the mutated LTRs leading into evolution. Thus, the mutants B7 and C6B

that always rebounded among all LPVs seems to indicate that G-to-A mutations anywhere in the viral LTR affect viral transfection, infection, and activation to PMA/I but can be counteracted with compensatory mutations however, mutations found specifically in TAR are the more difficult to compensate. Similarly, the other clones that rebounded including F8 and G9 have several mutations in the LTR including some TAR mutations making them hard to revert. Interestingly, the LPVs that rebounded in this experiment were also ones that have constantly shown signs of exhibiting a latency-like phenotype.

This project started off with 57 mutants (**Figure 4.1 and 4.2**), out of which 12 were shortlisted initially. In the various experiments conducted, we analyzed the 12 shortlisted LPVs and at the end of each experiment we discussed the clones that were leaning towards displaying a latency-like phenotype. Looking back, we can see that the same set of LPVs were seen to show latency-like phenotype across almost all the experiments with exception to a few that were disregarded during different experiments. By definition, the various properties displayed by a latent virus are low/no transcription that eventually leads to low infection, ability to respond to latency reversing agents, ability to revert from the state of latency when existing conditions change [32, 146]. Overall, the LPVs B7, C6b are proving our hypothesis that, G-to-A mutations introduced by A3 results in a latency-like phenotype. In addition, further trials will help in confirming if the LPVs D3, F8 and G9 are capable of reverting. As discussed before, these latter LPVs have mutations in the TAR region making them very difficult to acquire compensatory mutations.

6 CONCLUSION

HIV related literature has rarely addressed viral molecular latency as one of the possible latency causing mechanisms. We started with 57 clones in the library from which 12 of the clones were initially shortlisted. These 12 mutants were scrutinized and tested throughout this thesis. We finally conclude based on all the results obtained that some of these mutations have the ability to attain a latency-like phenotype that could be re-activated when treated with PMA/I. We have mainly focussed on various mutations in NF-κB, NFAT, and Sp-1 transcription factor binding sites, and TAR region that are responsible for less efficient viral transcription that was expected out of these LPVs. The reduced viral transcription was predicted to be due to loss of transcription factor binding mainly NF-κB, NFAT and Sp-1 by the in-silico experiment. Induction with PMA/I might compensate the latency inducing effect of a few of these binding site mutations by activating the NF-κB pathway of a few clones. The clones that reactivated were surprisingly consistent over all the various experiments performed. We have shown that mutations introduced by A3G/F in the HIV LTR region led to latency-like phenotype in this thesis satisfying our hypothesis. This thesis proposes a new pathway for HIV latency that will aid in future attempts at discovering new therapeutic strategy to tackle HIV infection.

7 REFERENCE

1. Barré-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 868-71.
2. Gallo, R.C. and L. Montagnier, *The discovery of HIV as the cause of AIDS.* N Engl J Med, 2003. **349**(24): p. 2283-5.
3. Roebuck, K.A. and M. Saifuddin, *Regulation of HIV-1 transcription.* Gene Expr, 1999. **8**(2): p. 67-84.
4. Masenga, S.K., et al., *HIV-Host Cell Interactions.* Cells, 2023. **12**(10).
5. Fanales-Belasio, E., et al., *HIV virology and pathogenetic mechanisms of infection: a brief overview.* Ann Ist Super Sanita, 2010. **46**(1): p. 5-14.
6. Campbell, E.M. and T.J. Hope, *HIV-1 capsid: the multifaceted key player in HIV-1 infection.* Nat Rev Microbiol, 2015. **13**(8): p. 471-83.
7. Vogt, s.V., *Retroviral Virions and Genomes. In: Coffin JM, Hughes SH, Varmus HE, editors. Retroviruses.* Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press;, 1997.
8. Stoltzfus, C.M., *Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication.* Adv Virus Res, 2009. **74**: p. 1-40.
9. Frankel, A.D. and J.A. Young, *HIV-1: fifteen proteins and an RNA.* Annu Rev Biochem, 1998. **67**: p. 1-25.
10. White, T.A., et al., *Molecular architectures of trimeric SIV and HIV-1 envelope glycoproteins on intact viruses: strain-dependent variation in quaternary structure.* PLoS Pathog, 2010. **6**(12): p. e1001249.
11. Arhel, N., *Revisiting HIV-1 uncoating.* Retrovirology, 2010. **7**: p. 96.
12. Kinoshita, S., et al., *The T cell activation factor NF-ATc positively regulates HIV-1 replication and gene expression in T cells.* Immunity, 1997. **6**(3): p. 235-44.
13. Nabel, G. and D. Baltimore, *An inducible transcription factor activates expression of human immunodeficiency virus in T cells.* Nature, 1987. **326**(6114): p. 711-3.
14. Liu, J., et al., *Specific NF-kappa B subunits act in concert with Tat to stimulate human immunodeficiency virus type 1 transcription.* J Virol, 1992. **66**(6): p. 3883-7.
15. Dingwall, C., et al., *HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure.* Embo j, 1990. **9**(12): p. 4145-53.
16. Rabbi, M.F., et al., *U5 region of the human immunodeficiency virus type 1 long terminal repeat contains TRE-like cAMP-responsive elements that bind both AP-1 and CREB/ATF proteins.* Virology, 1997. **233**(1): p. 235-45.

17. Darcis, G., et al., *Molecular Control of HIV and SIV Latency*, in *HIV-1 Latency*, G. Silvestri and M. Lichterfeld, Editors. 2018, Springer International Publishing: Cham. p. 1-22.
18. Board, N.L., et al., *Engaging innate immunity in HIV-1 cure strategies.* Nature Reviews Immunology, 2022. **22**(8): p. 499-512.
19. Van Lint, C., S. Bouchat, and A. Marcello, *HIV-1 transcription and latency: an update.* Retrovirology, 2013. **10**(1): p. 67.
20. Henderson, L.J., et al., *Advances toward Curing HIV-1 Infection in Tissue Reservoirs.* Journal of Virology, 2020. **94**(3): p. 10.1128/jvi.00375-19.
21. Borrell, M., et al., *High rates of long-term progression in HIV-1-positive elite controllers.* J Int AIDS Soc, 2021. **24**(2): p. e25675.
22. Allers, K., et al., *Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation.* Blood, 2011. **117**(10): p. 2791-9.
23. Gupta, R.K., et al., *HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation.* Nature, 2019. **568**(7751): p. 244-248.
24. Ding, J., Y. Liu, and Y. Lai, *Knowledge From London and Berlin: Finding Threads to a Functional HIV Cure.* Front Immunol, 2021. **12**: p. 688747.
25. Peterhoff, D., *New case of HIV cure: joined forces of haploidentical stem cells and HLA-mismatched cord blood.* Signal Transduction and Targeted Therapy, 2023. **8**(1): p. 241.
26. Clutter, C., *Conquering Latent HIV To End the AIDS Pandemic*, in *American Society for Microbiology*. 2020.
27. Ho, Y.C., et al., *Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure.* Cell, 2013. **155**(3): p. 540-51.
28. Avettand-Fènoël, V., et al., *Total HIV-1 DNA, a Marker of Viral Reservoir Dynamics with Clinical Implications.* Clinical Microbiology Reviews, 2016. **29**(4): p. 859-880.
29. Romani, B. and E. Allahbakhshi, *Underlying mechanisms of HIV-1 latency.* Virus Genes, 2017. **53**(3): p. 329-339.
30. Alexaki, A., Y. Liu, and B. Wigdahl, *Cellular reservoirs of HIV-1 and their role in viral persistence.* Curr HIV Res, 2008. **6**(5): p. 388-400.
31. Chaillon, A., et al., *HIV persists throughout deep tissues with repopulation from multiple anatomical sources.* The Journal of Clinical Investigation, 2020. **130**(4): p. 1699-1712.
32. Wong, J.K. and S.A. Yukl, *Tissue reservoirs of HIV.* Curr Opin HIV AIDS, 2016. **11**(4): p. 362-70.

33. Yukl, S.A., et al., *The distribution of HIV DNA and RNA in cell subsets differs in gut and blood of HIV-positive patients on ART: implications for viral persistence.* J Infect Dis, 2013. **208**(8): p. 1212-20.
34. Finzi, D., 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): p. 512-7.
35. Chun, T.W., et al., *In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency.* Nat Med, 1995. **1**(12): p. 1284-90.
36. Finzi, D., et al., *Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy.* Science, 1997. **278**(5341): p. 1295-300.
37. Siliciano, J.D., et al., *Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells.* Nature Medicine, 2003. **9**(6): p. 727-728.
38. Chomont, N., et al., *HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation.* Nat Med, 2009. **15**(8): p. 893-900.
39. Bosque, A., et al., *Homeostatic proliferation fails to efficiently reactivate HIV-1 latently infected central memory CD4+ T cells.* PLoS Pathog, 2011. **7**(10): p. e1002288.
40. Vandergeeten, C., et al., *Interleukin-7 promotes HIV persistence during antiretroviral therapy.* Blood, 2013. **121**(21): p. 4321-9.
41. Cohn, L.B., N. Chomont, and S.G. Deeks, *The Biology of the HIV-1 Latent Reservoir and Implications for Cure Strategies.* Cell Host Microbe, 2020. **27**(4): p. 519-530.
42. Koelsch, K.K., et al., *Dynamics of total, linear nonintegrated, and integrated HIV-1 DNA in vivo and in vitro.* J Infect Dis, 2008. **197**(3): p. 411-9.
43. Sloan, R.D. and M.A. Wainberg, *The role of unintegrated DNA in HIV infection.* Retrovirology, 2011. **8**: p. 52.
44. Lee, Y.M. and J.M. Coffin, *Efficient autointegration of avian retrovirus DNA in vitro.* J Virol, 1990. **64**(12): p. 5958-65.
45. Katz, R.A., et al., *Restriction endonuclease and nucleotide sequence analyses of molecularly cloned unintegrated avian tumor virus DNA: structure of large terminal repeats in circle junctions.* J Virol, 1982. **42**(1): p. 346-51.
46. Wiskerchen, M. and M.A. Muesing, *Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells.* J Virol, 1995. **69**(1): p. 376-86.
47. Kantor, B., et al., *Epigenetic activation of unintegrated HIV-1 genomes by gut-associated short chain fatty acids and its implications for HIV infection.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18786-91.

48. Sloan, R.D., et al., *Expression of Nef from unintegrated HIV-1 DNA downregulates cell surface CXCR4 and CCR5 on T-lymphocytes.* Retrovirology, 2010. **7**: p. 44.
49. Gillim-Ross, L., A. Cara, and M.E. Klotman, *Nef expressed from human immunodeficiency virus type 1 extrachromosomal DNA downregulates CD4 on primary CD4+ T lymphocytes: implications for integrase inhibitors.* J Gen Virol, 2005. **86**(Pt 3): p. 765-771.
50. Poon, B. and I.S. Chen, *Human immunodeficiency virus type 1 (HIV-1) Vpr enhances expression from unintegrated HIV-1 DNA.* J Virol, 2003. **77**(7): p. 3962-72.
51. Poon, B., M.A. Chang, and I.S. Chen, *Vpr is required for efficient Nef expression from unintegrated human immunodeficiency virus type 1 DNA.* J Virol, 2007. **81**(19): p. 10515-23.
52. Thierry, S., et al., *Integrase inhibitor reversal dynamics indicate unintegrated HIV-1 dna initiate de novo integration.* Retrovirology, 2015. **12**: p. 24.
53. Chun, T.W., et al., *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): p. 8869-73.
54. El Kharroubi, A., et al., *Transcriptional activation of the integrated chromatin-associated human immunodeficiency virus type 1 promoter.* Mol Cell Biol, 1998. **18**(5): p. 2535-44.
55. Verdin, E., P. Paras, Jr., and C. Van Lint, *Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation.* Embo j, 1993. **12**(8): p. 3249-59.
56. Rose, N.R. and R.J. Klose, *Understanding the relationship between DNA methylation and histone lysine methylation.* Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2014. **1839**(12): p. 1362-1372.
57. Eberharter, A. and P.B. Becker, *Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics.* EMBO Rep, 2002. **3**(3): p. 224-9.
58. Imai, K., H. Togami, and T. Okamoto, *Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294.* J Biol Chem, 2010. **285**(22): p. 16538-45.
59. Tripathy, M.K., et al., *H3K27 Demethylation at the Proviral Promoter Sensitizes Latent HIV to the Effects of Vorinostat in Ex Vivo Cultures of Resting CD4+ T Cells.* J Virol, 2015. **89**(16): p. 8392-405.
60. Kauder, S.E., et al., *Epigenetic Regulation of HIV-1 Latency by Cytosine Methylation.* PLOS Pathogens, 2009. **5**(6): p. e1000495.
61. Bird, A.P., *CpG-rich islands and the function of DNA methylation.* Nature, 1986. **321**(6067): p. 209-13.

62. Medvedeva, Y.A., et al., *Effects of cytosine methylation on transcription factor binding sites.* BMC Genomics, 2014. **15**(1): p. 119.
63. Weber, S., et al., *Epigenetic analysis of HIV-1 proviral genomes from infected individuals: predominance of unmethylated CpG's.* Virology, 2014. **449**: p. 181-9.
64. Lewinski, M.K., et al., *Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription.* J Virol, 2005. **79**(11): p. 6610-9.
65. Schröder, A.R., et al., *HIV-1 integration in the human genome favors active genes and local hotspots.* Cell, 2002. **110**(4): p. 521-9.
66. Lenasi, T., X. Contreras, and B.M. Peterlin, *Transcriptional interference antagonizes proviral gene expression to promote HIV latency.* Cell Host Microbe, 2008. **4**(2): p. 123-33.
67. Han, Y., et al., *Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough.* Cell Host Microbe, 2008. **4**(2): p. 134-46.
68. Shan, L., et al., *Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model.* J Virol, 2011. **85**(11): p. 5384-93.
69. Adelman, K. and J.T. Lis, *Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans.* Nat Rev Genet, 2012. **13**(10): p. 720-31.
70. Budhiraja, S. and A.P. Rice, *Reactivation of latent HIV: do all roads go through P-TEFb?* Future Virol, 2013. **8**(7).
71. Kumar, A., et al., *Epigenetic control of HIV-1 post integration latency: implications for therapy.* Clin Epigenetics, 2015. **7**: p. 103.
72. Alanazi, A., et al., *Targeting Tat-TAR RNA Interaction for HIV-1 Inhibition.* Viruses, 2021. **13**(10).
73. Khan, M.T., et al., *Structure-based analysis of the molecular recognitions between HIV-1 TAR-RNA and transcription factor nuclear factor-kappaB (NFkB).* Curr Top Med Chem, 2012. **12**(8): p. 814-27.
74. Gotora, P.T., R. van der Sluis, and M.E. Williams, *HIV-1 Tat amino acid residues that influence Tat-TAR binding affinity: a scoping review.* BMC Infect Dis, 2023. **23**(1): p. 164.
75. Barboric, M., et al., *NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II.* Mol Cell, 2001. **8**(2): p. 327-37.
76. Duverger, A., et al., *An AP-1 binding site in the enhancer/core element of the HIV-1 promoter controls the ability of HIV-1 to establish latent infection.* J Virol, 2013. **87**(4): p. 2264-77.
77. Abdel-Mohsen, M., et al., *Human Galectin-9 Is a Potent Mediator of HIV Transcription and Reactivation.* PLoS Pathog, 2016. **12**(6): p. e1005677.

78. Ganesh, L., et al., *The gene product Murr1 restricts HIV-1 replication in resting CD4+ lymphocytes.* Nature, 2003. **426**(6968): p. 853-857.
79. Tyagi, M., R.J. Pearson, and J. Karn, *Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction.* J Virol, 2010. **84**(13): p. 6425-37.
80. Williams, S.A. and W.C. Greene, *Regulation of HIV-1 latency by T-cell activation.* Cytokine, 2007. **39**(1): p. 63-74.
81. Khoury, G., et al., *The Molecular Biology of HIV Latency.* Adv Exp Med Biol, 2018. **1075**: p. 187-212.
82. Ajoge, H.O., et al., *Antiretroviral APOBEC3 cytidine deaminases alter HIV-1 provirus integration site profiles.* Nat Commun, 2023. **14**(1): p. 16.
83. Gaebler, C., et al., *Sequence Evaluation and Comparative Analysis of Novel Assays for Intact Proviral HIV-1 DNA.* J Virol, 2021. **95**(6).
84. Eriksson, S., et al., *Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies.* PLoS Pathog, 2013. **9**(2): p. e1003174.
85. Wilhelm, E., et al., *CTGC motifs within the HIV core promoter specify Tat-responsive pre-initiation complexes.* Retrovirology, 2012. **9**: p. 62.
86. Li, Y.L., et al., *The structural basis for HIV-1 Vif antagonism of human APOBEC3G.* Nature, 2023. **615**(7953): p. 728-733.
87. Ran, X.H., et al., *TRIM5α recruits HDAC1 to p50 and Sp1 and promotes H3K9 deacetylation at the HIV-1 LTR.* Nat Commun, 2023. **14**(1): p. 3343.
88. Stremlau, M., et al., *The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys.* Nature, 2004. **427**(6977): p. 848-53.
89. Neil, S.J., T. Zang, and P.D. Bieniasz, *Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu.* Nature, 2008. **451**(7177): p. 425-30.
90. Sharifi, H.J., et al., *Sulforaphane Reduces SAMHD1 Phosphorylation To Protect Macrophages from HIV-1 Infection.* J Virol, 2022. **96**(23): p. e0118722.
91. Yang, Z. and W.C. Greene, *A new activity for SAMHD1 in HIV restriction.* Nat Med, 2014. **20**(8): p. 808-9.
92. Nasr, N., et al., *HIV-1 infection of human macrophages directly induces viperin which inhibits viral production.* Blood, 2012. **120**(4): p. 778-88.
93. Staeheli, P. and O. Haller, *Human MX2/MxB: a Potent Interferon-Induced Postentry Inhibitor of Herpesviruses and HIV-1.* J Virol, 2018. **92**(24).
94. Yang, J., J. Hou, and M. Li, *Research on the influence of APOBEC family on the occurrence, diagnosis, and treatment of various tumors.* J Cancer Res Clin Oncol, 2023. **149**(1): p. 357-366.

95. Willems, L. and N.A. Gillet, *APOBEC3 Interference during Replication of Viral Genomes.* Viruses, 2015. **7**(6): p. 2999-3018.
96. Refsland, E.W., et al., *Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction.* Nucleic Acids Res, 2010. **38**(13): p. 4274-84.
97. Roberts, S.A., et al., *An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers.* Nature Genetics, 2013. **45**(9): p. 970-976.
98. Smith, H.C., et al., *Functions and regulation of the APOBEC family of proteins.* Semin Cell Dev Biol, 2012. **23**(3): p. 258-68.
99. Delviks-Frankenberry, K.A., B.A. Desimmie, and V.K. Pathak, *Structural Insights into APOBEC3-Mediated Lentiviral Restriction.* Viruses, 2020. **12**(6): p. 587.
100. Burdick, R.C., et al., *HIV-1 uncoats in the nucleus near sites of integration.* Proceedings of the National Academy of Sciences, 2020. **117**(10): p. 5486-5493.
101. Mangeat, B., et al., *Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts.* Nature, 2003. **424**(6944): p. 99-103.
102. Liddament, M.T., et al., *APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo.* Curr Biol, 2004. **14**(15): p. 1385-91.
103. Ulenga, N.K., et al., *The level of APOBEC3G (hA3G)-related G-to-A mutations does not correlate with viral load in HIV type 1-infected individuals.* AIDS Res Hum Retroviruses, 2008. **24**(10): p. 1285-90.
104. Harris, R.S. and M.T. Liddament, *Retroviral restriction by APOBEC proteins.* Nat Rev Immunol, 2004. **4**(11): p. 868-77.
105. Feng, Y., et al., *Suppression of APOBEC3-mediated restriction of HIV-1 by Vif.* Front Microbiol, 2014. **5**: p. 450.
106. Li, A., et al., *HBV triggers APOBEC2 expression through miR-122 regulation and affects the proliferation of liver cancer cells.* Int J Oncol, 2019. **55**(5): p. 1137-1148.
107. Covino, D.A., M.C. Gauzzi, and L. Fantuzzi, *Understanding the regulation of APOBEC3 expression: Current evidence and much to learn.* Journal of Leukocyte Biology, 2017. **103**(3): p. 433-444.
108. Rasmussen, T.A., M. Tolstrup, and O.S. Søgaard, *Reversal of Latency as Part of a Cure for HIV-1.* Trends Microbiol, 2016. **24**(2): p. 90-97.
109. Spivak, A.M. and V. Planelles, *Novel Latency Reversal Agents for HIV-1 Cure.* Annual Review of Medicine, 2018. **69**(1): p. 421-436.
110. Deeks, S.G., *HIV: Shock and kill.* Nature, 2012. **487**(7408): p. 439-40.

111. Shan, L., et al., *Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation.* Immunity, 2012. **36**(3): p. 491-501.
112. Jiang, G. and S. Dandekar, *Targeting NF-κB signaling with protein kinase C agonists as an emerging strategy for combating HIV latency.* AIDS Res Hum Retroviruses, 2015. **31**(1): p. 4-12.
113. Spivak, A.M. and V. Planelles, *HIV-1 Eradication: Early Trials (and Tribulations).* Trends in Molecular Medicine, 2016. **22**(1): p. 10-27.
114. Macedo, A.B., C.L. Novis, and A. Bosque, *Targeting Cellular and Tissue HIV Reservoirs With Toll-Like Receptor Agonists.* Front Immunol, 2019. **10**: p. 2450.
115. Doyon, G., et al., *Disulfiram reactivates latent HIV-1 expression through depletion of the phosphatase and tensin homolog.* Aids, 2013. **27**(2): p. F7-f11.
116. Archin, N.M., et al., *Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy.* Nature, 2012. **487**(7408): p. 482-485.
117. Søgaard, O.S., et al., *The Depsipeptide Romidepsin Reverses HIV-1 Latency In Vivo.* PLOS Pathogens, 2015. **11**(9): p. e1005142.
118. Margolis, D.M., *Histone deacetylase inhibitors and HIV latency.* Curr Opin HIV AIDS, 2011. **6**(1): p. 25-9.
119. DeChristopher, B.A., et al., *Designed, synthetically accessible bryostatin analogues potently induce activation of latent HIV reservoirs in vitro.* Nat Chem, 2012. **4**(9): p. 705-10.
120. Molitor, J.A., et al., *NF-kappa B: a family of inducible and differentially expressed enhancer-binding proteins in human T cells.* Proc Natl Acad Sci U S A, 1990. **87**(24): p. 10028-32.
121. Karn, J. and C.M. Stoltzfus, *Transcriptional and posttranscriptional regulation of HIV-1 gene expression.* Cold Spring Harb Perspect Med, 2012. **2**(2): p. a006916.
122. Bullen, C.K., et al., *New ex vivo approaches distinguish effective and ineffective single agents for reversing HIV-1 latency in vivo.* Nat Med, 2014. **20**(4): p. 425-9.
123. Laird, G.M., et al., *Ex vivo analysis identifies effective HIV-1 latency-reversing drug combinations.* J Clin Invest, 2015. **125**(5): p. 1901-12.
124. Mori, L. and S.T. Valente, *Cure and Long-Term Remission Strategies*, in *HIV Reservoirs: Methods and Protocols*, G. Poli, E. Vicenzi, and F. Romerio, Editors. 2022, Springer US: New York, NY. p. 391-428.
125. Ne, E., R.-J. Palstra, and T. Mahmoudi, *Chapter Six - Transcription: Insights From the HIV-1 Promoter*, in *International Review of Cell and Molecular Biology*, F. Loos, Editor. 2018, Academic Press. p. 191-243.
126. Van Duyne, R., et al., *Effect of mimetic CDK9 inhibitors on HIV-1-activated transcription.* J Mol Biol, 2013. **425**(4): p. 812-29.

127. Okamoto, M., et al., *Selective inhibition of HIV-1 replication by the CDK9 inhibitor FIT-039.* Antiviral Res, 2015. **123**: p. 1-4.
128. Lacombe, B., et al., *Specific Inhibition of HIV Infection by the Action of Spironolactone in T Cells.* J Virol, 2016. **90**(23): p. 10972-10980.
129. Vozzolo, L., et al., *Gyrase B inhibitor impairs HIV-1 replication by targeting Hsp90 and the capsid protein.* J Biol Chem, 2010. **285**(50): p. 39314-28.
130. Zhu, J., et al., *Reactivation of latent HIV-1 by inhibition of BRD4.* Cell Rep, 2012. **2**(4): p. 807-16.
131. Alamer, E., et al., *Epigenetic Suppression of HIV in Myeloid Cells by the BRD4-Selective Small Molecule Modulator ZL0580.* J Virol, 2020. **94**(11).
132. Schwarzer, R., A. Gramatica, and W.C. Greene, *Reduce and Control: A Combinatorial Strategy for Achieving Sustained HIV Remissions in the Absence of Antiretroviral Therapy.* Viruses, 2020. **12**(2).
133. Vansant, G., et al., *Block-And-Lock Strategies to Cure HIV Infection.* Viruses, 2020. **12**(1).
134. Stephenson, K.E., *Therapeutic vaccination for HIV: hopes and challenges.* Curr Opin HIV AIDS, 2018. **13**(5): p. 408-415.
135. Martinez-Navio, J.M., et al., *Long-Term Delivery of an Anti-SIV Monoclonal Antibody With AAV.* Front Immunol, 2020. **11**: p. 449.
136. Steinhaus, R., P.N. Robinson, and D. Seelow, *FABIAN-variant: predicting the effects of DNA variants on transcription factor binding.* Nucleic Acids Res, 2022. **50**(W1): p. W322-w329.
137. Dufour, C., et al., *The multifaceted nature of HIV latency.* J Clin Invest, 2020. **130**(7): p. 3381-3390.
138. Lian, X., et al., *Progressive transformation of the HIV-1 reservoir cell profile over two decades of antiviral therapy.* Cell Host Microbe, 2023. **31**(1): p. 83-96.e5.
139. Chen, J., et al., *The reservoir of latent HIV.* Front Cell Infect Microbiol, 2022. **12**: p. 945956.
140. Mbonye, U. and J. Karn, *The Molecular Basis for Human Immunodeficiency Virus Latency.* Annu Rev Virol, 2017. **4**(1): p. 261-285.
141. O'Brien, M.J., K. Gurdziel, and A. Ansari, *Global Run-On sequencing to measure nascent transcription in Saccharomyces cerevisiae.* Methods, 2023. **217**: p. 18-26.
142. Chen, Q.W., et al., *LncRNA BC promotes lung adenocarcinoma progression by modulating IMPAD1 alternative splicing.* Clin Transl Med, 2023. **13**(1): p. e1129.
143. Mingyan, Y., L. Xinyong, and E. De Clercq, *NF-kappaB: the inducible factors of HIV-1 transcription and their inhibitors.* Mini Rev Med Chem, 2009. **9**(1): p. 60-9.

144. Kim, T.H. and J. Dekker, *ChIP-Quantitative Polymerase Chain Reaction (ChIP-qPCR).* Cold Spring Harb Protoc, 2018. **2018**(5).
145. Lu, F., et al., *Epigenetic Landscape of HIV-1 Infection in Primary Human Macrophage.* J Virol, 2022. **96**(7): p. e0016222.
146. Lau, C.Y., M.A. Adan, and F. Maldarelli, *Why the HIV Reservoir Never Runs Dry: Clonal Expansion and the Characteristics of HIV-Infected Cells Challenge Strategies to Cure and Control HIV Infection.* Viruses, 2021. **13**(12).
147. Lusic, M. and R.F. Siliciano, *Nuclear landscape of HIV-1 infection and integration.* Nat Rev Microbiol, 2017. **15**(2): p. 69-82.
148. Liszewski, M.K., J.J. Yu, and U. O'Doherty, *Detecting HIV-1 integration by repetitive-sampling Alu-gag PCR.* Methods, 2009. **47**(4): p. 254-60.
149. Brady, T., et al., *Quantitation of HIV DNA integration: effects of differential integration site distributions on Alu-PCR assays.* J Virol Methods, 2013. **189**(1): p. 53-7.
150. Swan, D.A., et al., *Evolution during primary HIV infection does not require adaptive immune selection.* Proc Natl Acad Sci U S A, 2022. **119**(7).
151. Rojas Sánchez, P., et al., *Impact of Clinical Parameters in the Intrahost Evolution of HIV-1 Subtype B in Pediatric Patients: A Machine Learning Approach.* Genome Biol Evol, 2017. **9**(10): p. 2715-2726.
152. Liu, R.D., et al., *Mechanism and factors that control HIV-1 transcription and latency activation.* J Zhejiang Univ Sci B, 2014. **15**(5): p. 455-65.
153. Pasternak, A.O., et al., *Editorial: The relevance of molecular mechanisms in HIV-1 latency and reactivation from latency.* Front Cell Infect Microbiol, 2023. **13**: p. 1190867.
154. Tarasova, O., et al., *Identification of Molecular Mechanisms Involved in Viral Infection Progression Based on Text Mining: Case Study for HIV Infection.* Int J Mol Sci, 2023. **24**(2).
155. Sabin, C.A. and J.D. Lundgren, *The natural history of HIV infection.* Curr Opin HIV AIDS, 2013. **8**(4): p. 311-7.
156. Rice, A.P., *Unexpected Mutations in HIV-1 That Confer Resistance to the Tat Inhibitor Didehydro-Cortistatin A.* mBio, 2019. **10**(4).
157. Bannwarth, S. and A. Gatignol, *HIV-1 TAR RNA: the target of molecular interactions between the virus and its host.* Curr HIV Res, 2005. **3**(1): p. 61-71.
158. Das, A.T., et al., *Opening of the TAR hairpin in the HIV-1 genome causes aberrant RNA dimerization and packaging.* Retrovirology, 2012. **9**: p. 59.

8 APPENDIX

8.1 Creation of Latency-Prone Viruses by Cindy Lam

The text below is reproduced from thesis written by Cindy Lam.

Cell culture

HEK 293T cells were maintained in complete DMEM. Cells were propagated in T-75 culture flasks (Corning Inc., Tewksbury, Massachusetts, USA) at 37 °C in a 5% CO2 incubator (Forma Series II Water Jacket CO2 Incubator) (Thermo Scientific Inc., Burlington, Ontario, Canada). Initial Plasmids The viral vector encoding single-round HIV-1 (pNL4.3-Δenv-eGFP) was obtained through the NIH AIDS Reagent Program (Catalog #11100). The Env gene of this vector has been disrupted with an enhanced green fluorescent protein (eGFP) reporter gene and rendered nonfunctional. Our lab further manipulated this vector by inactivating the Vif gene via two stop codons introduced by site-directed mutagenesis (SDM), herein referred to as pNL4-3-ΔenveGFPΔVif. VSV-G envelope plasmid used to pseudotype single-round HIV-1 infections were obtained through Addgene (Cambridge, Massachusetts, USA). The human APOBEC2, APOBEC3G, and APOBEC3F expression plasmids have been described previously (Langlois et al., 2009; Takeda et al., 2008; Bélanger et al., 2013).

Transfection and Infections

HEK 293T cells were used to produce the infectious VSV-G-pseudotyped HIV-1 viral particles with packaged APOBEC3 proteins in co-transfections. 3.0 x 105 cells were seeded per well of a 6-well plate in 2 mL of complete DMEM and incubated for approximately 24 hours until they reached 80% confluency. Prior to transfection, the media was replaced with 3 mL of fresh complete DMEM. Co-transfections were performed with 500ng of pNL4-3-ΔenveGFPΔVif plasmid, 200ng of VSV-G and 150ng of either APOBEC3G or APOBEC3F plasmid. 117 Expression vectors were mixed with GeneJuice transfection reagent (EMD Millipore Corp., Billerica, Massachusetts, USA) and used to transfect the HEK 293T producer cells according to the manufacturer's specifications. The cells were then incubated for 48 hours to allow for production of infectious virus and packaged APOBEC3 protein. 24 hours prior to infection, 3.0 x

105 HEK 293T cells were seeded per well of a 6 well plate in 2 mL of DMEM and incubated until they reached 50-60% confluency. Prior to infection, the media was replaced with 3 mL of fresh DMEM plus polybrene (Sigma-Aldrich Canada Co., Oakville, Ontario, Canada) at a final concentration of 8 μg/mL. 48 hours post-transfection, virus-containing supernatants from HEK 293T producer cells were collected and cleared by centrifugation at 2000rpm for 10 minutes in a Sorvall ST40R centrifuge (Thermo Scientific). The HEK 293T target cells were infected with equal volumes of the viral supernatants by spinoculation at 900xg for 1 hour and then incubated at 37 °C. At 48 hours post-infection, infectivity was analyzed by monitoring expression of the integrated eGFP reporter gene by flow cytometry using the CyAN ADP9 flow cytometer (Beckman Coulter, Inc., Brea, California, USA).

PCR

In addition to flow cytometry analysis on the infected HEK 293T target cells, a portion of the cells was collected and lysed for genomic DNA (gDNA) extraction. The gDNA was extracted using the Wizard Genomic DNA Purification Kit (Promega Corp., Madison, Wisconsin, USA) according to manufacturer's specifications. Following gDNA extraction, a semi-nested polymerase chain reaction (PCR) was performed on the gDNA to specifically amplify both the 5' and 3' LTRs of the integrated proviruses. All PCR reactions were performed in a total 118 reaction volume of 50 μL with (final) 1U PrimeSTAR HS DNA polymerase (TaKaRa), 1X PrimeSTAR buffer (Mg2+), 200 μM dNTPs, 200 nM of each forward and reverse primers, and 10 ng of gDNA template for the first round (2 μL of a 1:25 dilution of the 1st round PCR products for the second round). The primers used to specifically amplify the 5' LTR were HIV LTR FWD deg 2 and HIV LTR REV 1 for the first round, and HIV LTR FWD deg 2 (and HIV LTR REV 3 deg for the second round (Lam et al, Table A1). The primers used to specifically amplify the 3' LTR were HIV 3' LTR FWD 1 and HIV 3' LTR REV deg for the first round, and HIV 3' LTR FWD 2 deg and HIV 3' LTR REV deg for the second round (Lam et al, Table A1). The second-round primers for both LTRs were designed to include restriction endonuclease sites (AseI and NheI) for downstream cloning applications. All primers were ordered from SigmaAldrich. The cycling conditions were as follows: an initial denaturation at 98 °C for 1 minute, followed by 32 cycles of denaturation at 98 °C for 10 seconds, annealing at 56 °C for 15

seconds (58 °C for 30 seconds for the second round), elongation at 72 °C for 1.5 minutes (1 minute for the second round), and a final elongation at 72 °C for 2 minutes. All PCRs were performed in an Eppendorf mastercycler Pro S series thermocycler.

Cloning

PCR products were resolved on 1% (w/v) agarose gels containing 1% ethidium bromide and visualized under UV light using the AlphaImager MINI Gel Imaging System (Alpha Innotech Corp., San Jose, California, USA) and accompanying software. The correct-sized bands (639 bp prior to digestion) were excised and gel purified using the Wizard SV Gel and PCR CleanUp System (Promega) according to manufacturer's specifications. The purified PCR products were then digested with the enzymes AseI and NheI (New England Biolabs Ltd. (NEB), Whitby, 119 Ontario, Canada) for 6 hours at 37 °C, column purified, and phosphorylated with T4 polynucleotide kinase (NEB) according to manufacturer's specifications. The peGFP-C3 expression vector was digested with the same enzymes as the PCR insert and dephosphorylated using Antarctic phosphatase (NEB) according to manufacturer's specifications prior to gel purification. Using T4 DNA ligase (NEB) according to manufacturer's specifications, the digested and phosphorylated PCR inserts were ligated in the place of the Cytomegalovirus (CMV) promoter upstream of the eGFP reporter gene of the similarly digested and dephosphorylated peGFP-C3 vector at varying ligation ratios (commonly 1:1, 1:3, and 1:5 vector:insert ratios). The ligations were performed in a total reaction volume of 10 µL and incubated at 16 °C or 4 °C overnight. The next day, the ligation reactions were mixed with 100 µL of DH5α competent E. coli cells, incubated on ice for 30 minutes prior to heat-shocking in a 42 °C water bath for 1.5 minutes and incubated on ice for 3 minutes. 900 µL of Luria Bertani (LB)-only broth was added to each of the transformation tubes and grown for at least one hour with shaking at 37 °C. The transformations were lightly pelleted at 1000 RPM for 5 minutes in a Sorvall Legend Micro 17 microcentrifuge (Thermo Scientific), resuspended in the remaining broth after decanting, and plated onto LB-Agar plates supplemented with 30 mg/mL kanamycin (KAN) for selection and incubated overnight at 37 °C. 120 Figure 21 Initial Cloning Strategy for LTR-eGFP Reporter Constructs

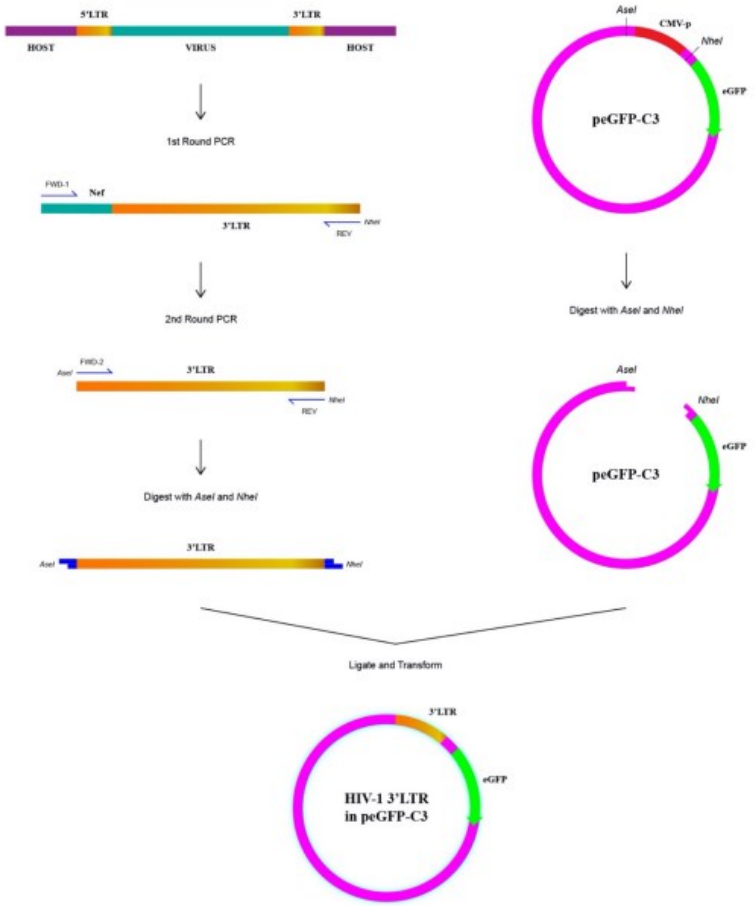


Figure 8.1a
Initial Cloning Strategy for LTR-eGFP Reporter Constructs

High Resolution Melt (HRM)

Individual colonies from the transformations were diluted in 250 µL of sterile water in 96- well plates. 8 µL of each colony dilution was added to another 96-well plate containing (final) 1X MeltDoctor HRM Master Mix (AmpliTaq Gold 360 DNA Polymerase, MeltDoctor HRM Dye, dNTP blend, Mg2+, buffer) (Applied Biosystems, Life Technologies Corp., Burlington, Ontario, Canada), 0.3 µM of each Seq peGFP-C3 LTR FWD and Seq peGFP-C3 LTR REV primers (Lam et al, Table A1), and topped up to 20 µL with sterile water. The reaction plate was briefly centrifuged at 2000 RPM for 2 minutes in a Sorvall centrifuge to bring all contents to the bottom of the plate. The cycling conditions for the real-time PCR (qPCR) reaction were as follows: an initial denaturation at 94 °C for 2 minutes, followed by 40 cycles of denaturation at 95 °C for 30 seconds, annealing at 55 °C for 45 seconds, and extension at 72 °C for 1 minute. Following the qPCR reaction, the samples underwent a melting curve analysis whereby each amplicon was heated up from 72 °C to 95 °C and fluorescence of the intercalated dye was measured in increments of 0.025 °C. The melting curves were aligned and samples that appeared to the left of the unmutated control curves were selected for Sanger sequencing to confirm G-to-A mutations. All HRM reactions were performed in a ViiA 7 real-time PCR system (Applied Biosystems).

Sequencing

The wells of the 96-well plate that contained the colony dilutions corresponding to the samples of interest from the HRM reaction were resuspended and each added to 1.5 mL of LB broth supplemented with kanamycin (minipreps). The minipreps were grown overnight for approximately 16-20 hours at 37 °C in a shaking incubator. The next day, the minipreps were processed using the Wizard Plus SV Minipreps DNA Purification System (Promega) according to manufacturer's specifications. 8 µL of each purified prep was sent out to the Nanuq Sequencing Facility at McGill University and Genome Quebec Innovation Centre for Sanger sequencing along with 10 µL of the forward and reverse sequencing primers, Seq peGFP-C3 LTR FWD and Seq peGFP-C3 LTR REV, respectively at 5.0 µM (Table A1). Upon receiving the sequencing results back from the facility, the sequences were aligned using Sequencher 4.10.1 software (Gene Codes Corp., Ann Arbor, Minnesota, USA) and analyzed for G-to-A mutations in an APOBEC3 deamination

context using the wildtype (wt) pNL4-3- Δenv-eGFPΔVif LTR (5' or 3' as applicable) sequence as a reference.

Functional Assays with LTR-eGFP Reporter Constructs

Uniquely-mutated clones as analyzed by sequencing were selected to generate a library of APOBEC3-mutated LTR-eGFP reporter constructs. HEK 293T cells were seeded at 1.0 x 105 cells per well of a 24-well plate in 0.5 mL of complete DMEM 24 hours prior to transfection until they reached approximately 80% confluency. Prior to transfection, the media was replaced with 1 mL of fresh DMEM. The cells were transfected with 100 ng of each clone alone as well as co-transfected with 100 ng of the tat expression plasmid using GeneJuice transfection reagent according to manufacturer's specifications. 48 hours post-transfection, the cells were harvested, and the functionality of the mutated promoters was assessed by measuring the expression and fluorescence intensity of the eGFP reporter gene using flow cytometry on a CyAN ADP9 flow cytometer.

HIV-1 Vector Backbone

A fragment of the viral vector pNL4-3-Δenv-eGFPΔVif was PCR amplified using (final) 2.5U Bestaq DNA polymerase (Applied Biological Materials Inc. (abm), Richmond, British Columbia, Canada), 1X PCR buffer (Mg2+), 200 µM dNTP mix, 200 nM of each of the primers pNL4-3 XhoI 3' LTR FWD and pNL4-3 NaeI-SacII 3' LTR REV (Table A1), 20 ng of template, and topped up to 50 µL with sterile water. The reverse primer incorporated an additional restriction endonuclease site (SacII) and spacer within the viral sequence for downstream cloning. The PCR conditions were as follows: an initial denaturation at 94 °C for 3 minutes, followed by 30 cycles of denaturation at 94 °C for 10 seconds, annealing at 58 °C for 30 seconds, extension at 72 °C for 30 seconds, and a final extension at 72 °C for 5 minutes. The amplicon was then digested with the restriction enzymes XhoI and NaeI (NEB), phosphorylated, and ligated into another pNL4-3-Δenv-eGFPΔVif vector previously digested with the same restriction enzymes and dephosphorylated. The ligations were transformed on LB-Agar plates supplemented with 100 mg/mL ampicillin (AMP) and minipreps were grown in LB+AMP broth and processed as previously described. This intermediate vector was further digested with the restriction enzymes SacII and NaeI (NEB) and

dephosphorylated in order to create the HIV-1 vector backbone in which the heavily mutated LTRs from the eGFP reporter library would be reconstituted. This HIV-1 vector backbone was sent for Sanger sequencing to confirm successful ligation of the intermediate fragment using the primer Seq HIV int-mut 3' LTR FWD.

Heavily Mutated LTR Inserts

The heavily mutated LTRs from the LTR-eGFP reporter library were PCR amplified using the same PCR reagents and conditions as above with the exception of the primers used (peGFPC3 SacII 3' LTR FWD deg and peGFP-C3 NaeI 3' LTR REV deg (Lam et al, Table A1)) and the extension time during cycles (45 seconds). The amplicons were then digested with the restriction enzymes SacII and NaeI, phosphorylated, and ligated into the HIV-1 vector backbone generated previously. The clonings were processed as described previously and successful reconstitution was confirmed through Sanger sequencing with the primers Seq HIV mut 3' LTR FWD and Seq HIV mut 3' LTR REV (Lam et al, Table A1) using each heavily mutated LTR from the eGFP reporter constructs as references for the reconstituted clones.

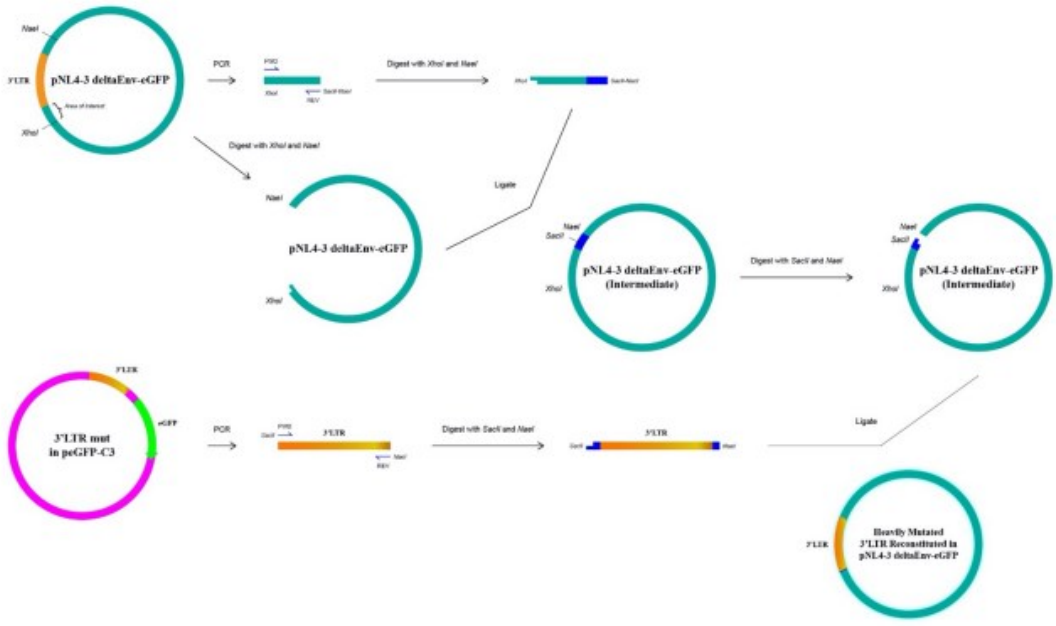


Figure 8.1b

Cloning Strategy for Reconstituted Latency-Prone Viral Plasmids and Heavily Mutated Clones.

8.2 CXCR4 LPV CLONING by Matthew Greig

The text below is reproduced from thesis written by Matthew Greig.

In order to move into a system that allows for multiple rounds of infections, our mutated LTRs needed to be cloned out of a non-replicative pNL4.3-Δenv-Δvif plasmid and into a fully replicative pNL4 CXCR4-tropic or pNL4 CCR5-tropic plasmid. A CXCR4 plasmid backbone was chosen as the first option for cloning because the majority of the viral infection cycle takes place in a CXCR4-tropic configuration where infectivity levels and disease progression markers are very high (194–196). We also saw higher preliminary levels of infection by the CXCR4-tropic pNL4 compared to our CCR5-tropic virus. To accomplish this transition, we used a simple cut and paste cloning strategy using a NcoI cut site downstream of the 3' end of the 3' LTR and an XhoI cut site upstream of the 5' end of the 3' LTR (within Nef). Mutated 3' LTRs were cut out of pNL4.3-Δenv-Δvif plasmids with 46 20U of NcoI-HF and 20U of XhoI. Products were electrophoresed on a 0.6% agarose gel and gel purified using the E.Z.N.A. Gel Extraction Kit (Omega BioTek D2500-02). Purified inserts were cloned into a pNL4.3 CXCR4-tropic digested backbone at a 5:1 ratio using T4 DNA Ligase (NEB #M0202S) at 16 oC overnight. Transformations of products were performed in Stbl-2 Ecoli cells using LB-Ampicillin plates. Colonies were picked and viral miniprep stocks generated as stated above under "latency-prone virus plasmid creation".

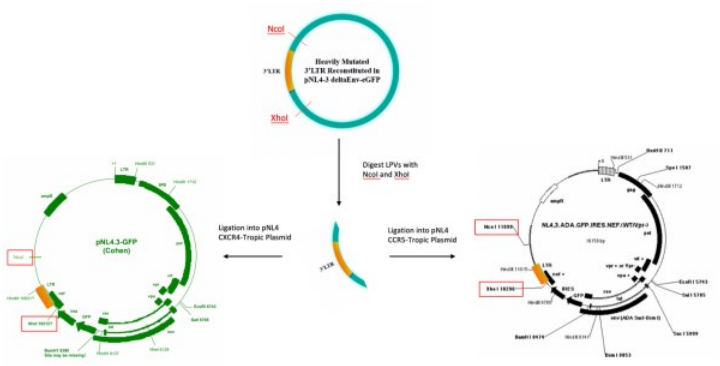


Figure 8.2

Cloning of LTRs into Replicative HIV-1 Backbone. LTRs of LPVs were cloned from a non-replicate pNL4- Δenv-Δvif plasmid into a fully replicative, CXCR4-tropic or CCR5-tropic NL4 plasmid. XhoI in the Nef and NcoI in the plasmid were used in the excision and ligation.

8.3 Similar Mutations Are Found in HIV-1 Patient Sequences in Los Alamos Database

The text below is reproduced from thesis written by Cindy Lam.

To lend further support to the mutations discovered in our in vitro experiments, the mutations in the 5' and 3' LTRs of the most prevalent subtypes and circulating recombinant forms (CRFs) in HIV-1-infected patient sequences recovered from the Los Alamos database were compared to those of our heavily mutated in vitro clones (Figure 4.8). The frequency of mutation and the transcription factor binding sites affected varied amongst the different subtypes, CRFs, and experimental clones. However, similar positions in transcription factor binding sites were mutated in the patient sequences as in our experimental clones, demonstrating that the mutations and transcription factor binding sites targeted in our experimental assays are relevant and also appear in nature.

To aid in the identification of potential mutations that could generate a latent phenotype, the mean eGFP fluorescence intensity for each position mutated was plotted for each clone. Separate box plots were generated for the A3G (Figure 4.9, Lam C thesis 2019) and A3F (Figure 4.10, Lam C thesis 2019) clone sets. Positions that are mutated and consistently generate a lowered MFI are desirable for the identification of potential latency-inducing mutations. Mutated positions that consistently generated an MFI as one of the criteria for selecting potential mutations from the A3G clone set. Because mutations did not modulate promoter activity as drastically in the A3F clone set, the threshold and criteria for selecting potential mutations from this set was increased to mutated positions that consistently generated an MFI ≤ 90.

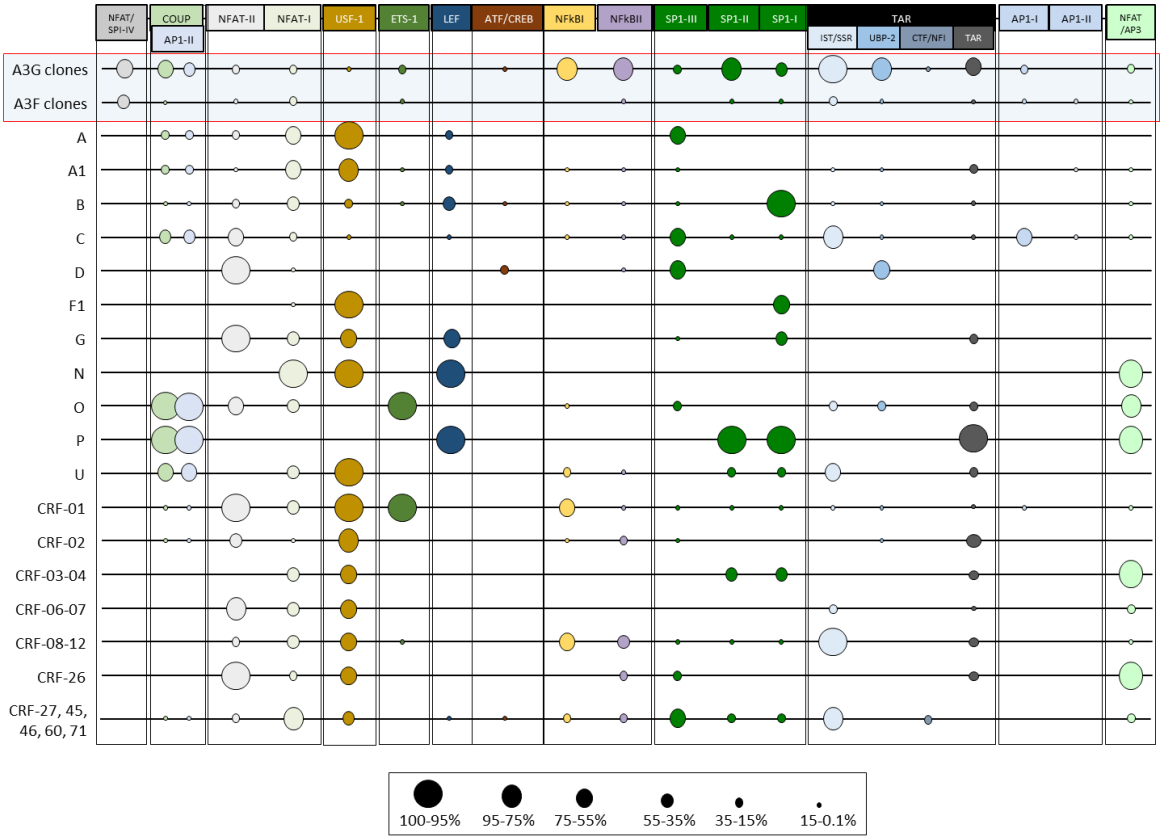


Figure 8.3 A3 mutations in patient sequences

8.4 Propagation of Antibodies for p24 ELISA

Hybridoma cells were thawed and centrifuged at 300xg for 5 min to remove freezing medium. Cells were resuspended in 10mL complete RPMI and cultured in T75 flask (Sigma-Aldrich #CLS430641, Oakville, ON, Canada). When confluent, cells were centrifuged at 300xg for 5 min in a 50mL tube, washed with 1X PBS+5mM EDTA, resuspended in 20mL complete RPMI media and placed in T175 flask (Sigma-Aldrich #CLS431080, Oakville, ON, Canada). This expansion was repeated every 2 days until 8 T175 confluent flasks were obtained. Once enough flasks were obtained, cells were collected, centrifuged at 300xg for 5 min, and washed with 1X PBS+ 5mM EDTA. Cells were then resuspended in 37oC warmed CD Hybridoma Media (Thermo Fisher #11279023, Waltham, MA, USA) supplemented with100U/mL penicillin and 100ug/mL streptomycin (P/S) and 8mM L-glutamine. Cells were placed in T175 flasks containing 50mL of complete media each. Once confluent, another 50mL of warmed complete hybridoma media was added to each flask before flasks were allowed to grow out for 2 weeks.

After 2 weeks cells and supernatant were collected and centrifuged in 50mL tubes at 3500rpm for 1-2 hours at 4oC. While cellular debris was pelleting from supernatant, column was prepared as seen in Figure X with 5mL of Protein G Agarose Beads (Exalpha #X1198, Shirley, MA, USA). The 50% slurry of beads was transferred to a centrifuge column and then centrifuged at 500xg for 5 min to collect beads before the solution was pipetted out and replaced with borate buffer (38.1g sodium borate decahydrate in 2L of Mili-Q water, brought to pH 8 with 36g of boric acid). Post spin supernatant was extracted, 0.2um filtered and combined with borate buffer to a ratio of 1L hybridoma supernatant: 2L borate buffer. Once supernatant + borate solution was ready, the bottom of the centrifuge column containing 127 the agarose beads and borate buffer was broken, and borate buffer was allowed to flow out into a waste beaker. 100mL of borate buffer was then added and allowed to flow through. After borate buffer finished flowing, the waste beaker was replaced with a 2L collection beaker, and all supernatant + borate solution was allowed to flow through the column, followed by 100mL of borate buffer to wash column (the last 100mL of borate buffer can go into the waste beaker). When only 1mL of buffer remained in the column, a 50mL tube containing 3mL of

Neutralization buffer (242.28g of Tris base in 1L sterile Mili-Q water, pH 9.5) was placed underneath column. 12mL of Elution buffer (8.92g citric acid, 1.04g sodium citrate dehydrate, 1L sterile Mili-Q water, pH 2.6) was placed into the top of the column and allowed to flow through the column and into the neutralization tube, followed by an additional 12mL of elution buffer. Neutralization tube was removed after flow through was complete, inverted gently to mix, and kept at 4oC as 'Elution #1'. 50mL of borate buffer was then run through into the waste beaker to wash the column. When borate wash was near completion

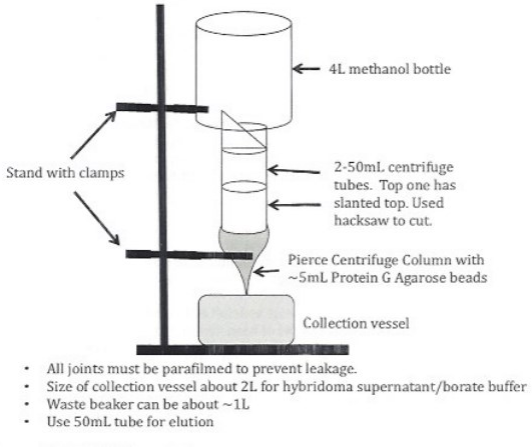


Figure 8.4 Column Setup for Antibody Purification

8.5 Primers and Probes

Primer and Probe Sequences

Primer/Probe Name	**5'-3' Sequence**
Lambda R Rev 2	GTTTCGCTTACGTGGCATCAGACGG
Alu1 Primer	TCCCAGCTACTGGGGAGGCTGAGG
Alu2 Primer	GCCTCCCAAAGTGCTGGGATTACAG
LTR R REGION Probe	5YAKYEI-CTTTATTGAGGCT+TAAG+C+AG+T+G+GGT-3IABKFQ (IDT)
Lambda R U5 Rev 1	AGTTTCGCTTACGTGGCATCAGACGGGCACACACTACTTTGAGCAC
Late U3 For	GCTACATATAAGCAGCTGCTTTTTGCCTGTAC
Actin Probe	56TAM-CCCTGTACGCCTCTGGCCGTACCACTGGC-3IABRQ
Actin 1	CATGTACGTTGCTATCCAGGC
Actin 2	CTCCTTAATGTCACGCACGAT

9 CV

Mrudhula Madapuji Srinivasan

EDUCATION

March 2021-Current

- MSc in Biochemistry Microbiology and Immunology
- University of Ottawa, Ottawa, ON

August 2018- July 2020

- M.Sc in Immunology
- Amity University, Noida, UP, India
- CGPA 7.58

June 2015- March 2018

- B.Sc. Zoology
- PSG College of Arts and Science, Coimbatore, TN, India
- CGPA 7.9

SCHOLARSHIPS

- Admission scholarship 2021
- Admission Scholarship 2022

RESEARCH EXPERIENCE

March 2021-Current

- MSc Candidate
- Dr. Marc-Andre Langlois Lab, University of Ottawa, Department of Biochemistry, Microbiology, and Immunology
- MSc Project: Molecular mechanisms and host factors involved in HIV-1 latency.
- Laboratory techniques used included viral transfections and infections, tissue culture work, PCR, qPCR, ELISAs, cellular staining, flow cytometry and Bioinformatics.

January 2020 – March 2020

- Vising student researcher to finish my dissertation at The University of Ottawa
- Dr. Subash Sad lab, University of Ottawa, Department of Biochemistry, Microbiology, and Immunology
- Topic- FoxO3a and Il-10 regulate inflammation and homeostasis.
- Laboratory techniques used included cell/tissue culture work, bacterial infection, Elisa, Western blotting.

June 2019-August 2019

- Summer internship at Indian Agricultural research institute (IARI), New Delhi, India
- Dr. Subramanian Department of Entomology, IARI, New Delhi.
- Topic- ROLE OF GUT MICROBES VIS-À-VIS BACTERIAL DISEASE INFECTION IN HONEYBEE *Apis cerana indica*
- Laboratory techniques used included isolation and comparison of gut microbiome DNA from diseased A. cerana indica honeybee larvae, PCR, bacterial culture and bioinformatics.

POSTERS AND PRESENTATIONS

POSTER

Srinivasan MM, Dankar S, Greig M, Lam C, Langlois MA. Effects of APOBEC3 mutations on HIV-1 LTR promoter on viral transcription and induction by Tat, 2022, Ottawa, ON.

Presented at

- BMI symposium 2022 held at Montebello organised by University of Ottawa.
- Antibody Diversification and DNA deaminases in Immunity and Cancer, held at Quebec City. I was also a part of the organising committee.

ORAL PRESENTATION

Srinivasan MM, Dankar S, Greig M, Lam C, Langlois MA. Molecular mechanisms and host factors involved in HIV-1 latency.

Presented at

- The university of Ottawa BMI day 2022
- The university of Ottawa BMI day 2023

PROFESSIONAL DEVELOPMENT

- Webinar by Michael Reth: Ramos B cell engineering for the evaluation of the humoral anti-SARS-CoV-2 immunity.
- Webinar by Elizabeth Mann and Madhvi Menon: Longitudinal immune profiling reveals distinct features of COVID-19 pathogenesis.
- Webinar by Sacha Gnjatic: Inflammatory cytokines and SARS-CoV-2 antibody responses predict survival and severity in COVID-19Two-day workshop- Flow cytometry.
- One-day workshop: "Molecular and Immuno techniques".
- One-day workshop: Honeybee-keeping
- One-day workshop: Modern trends in biology with a focus on Nano biotechnology
- One-day workshop: "Workshop in bioinformatics"
- Two days National workshop: "Intellectual Property Rights"
- One day Lecture-cum-workshop on the Role of Search engines in the field of Zoology.