EFFECTS OF APOBEC3-INDUCED MUTATIONS ON HIV-1 EXPRESSION AND LATENCY

Cindy Lam

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
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

© Cindy Lam, Ottawa, Canada, 2019

ABSTRACT

One of the greatest barriers to curing HIV-1 is the ability of the virus to establish a state of latent infection within infected cells. During viral latency, the virus lies dormant in the form of an integrated, replication-competent provirus within the infected cell. In this state, the virus is undetectable by the host immune system and is unaffected by antiretroviral drugs due to extremely weak or null transcription. These viruses, however, can be induced to produce infectious virus later on and propagate the infection as well as reseed the latent reservoir. The factors that lead to the establishment and maintenance of HIV-1 latency are not all known. Current latency reversal methods are unable to effectively purge the latent reservoir in HIV-1-infected patients, which begs the question whether there are populations of latently-infected cells that are not being targeted by present therapeutics.

One of our first lines of defense against retroviral infection is the APOBEC3 family of cytidine deaminases. These enzymes restrict retroviral infection mainly through hypermutation of the proviral DNA, leading to inactivation of the virus. However, it has been shown that low levels of mutations do not necessarily inactivate the virus and can sometimes be beneficial for the virus by increasing genetic diversity and viral fitness. Sublethal mutagenesis has been studied on coding regions of the virus, however, their effects on regulatory regions of the virus such as the LTR promoter have been scarcely explored. My project aimed to examine the effects of APOBEC3-induced mutations on the activity of the HIV-1 LTR promoter and investigate whether these mutations could be generating a subset of latently-infected cells.

A library of HIV-1 clones with A3G- or A3F-mutated LTRs was generated. We discovered that the 5' LTR is not mutated during the first round of HIV-1 infection in our system while the 3' LTR accumulates mutations. A second round of infection is required for

the mutations in the 3' LTR to be copied over to the 5' LTR and influence promoter activity. The mutations generated a range of effects on the promoter, with some clones being completely inactivated and no longer responsive to Tat or PMA/Ionomycin induction, while other clones were still highly infectious. The clones of interest were the ones in between this spectrum that were weakly-infectious (≤0.25%) prior to stimulation but were able to be induced above a given threshold (≥10-fold). We denoted these clones latency-prone viruses (LPVs) and were able to identify 10 of these clones within our library. We also explored the effects of individual mutations on the promoter and although these clones were highly infectious, we identified 8 positions of interest that led to weaker infectivity and fluorescence when mutated. The mutations in our library were found to hit important transcription factor binding sites and were also identified in a bioinformatics search of HIV-1-infected patient sequences, which confirms the relevance of our *in vitro*-identified mutations in a physiological setting.

This is the first investigation and evidence of the APOBEC3 proteins contributing to the generation of a subset of latent viruses and constitutes an important contribution to the understanding of HIV-1 latency and its reversal. This previously uncharacterized pool of latently-infected cells could explain why current therapies are ineffective as they do not target these molecular modifications but rather focus on reversing epigenetics and reactivation of the virus. These newly-discovered latency-prone viruses would be a useful tool in testing reactivation drugs and establishes the need to develop novel antiretrovirals that will target a broader spectrum of latently-infected cells. This project has effectively illustrated the dual role of this host-encoded restriction factor and provided further insight into HIV-1 latency, the major hurdle towards a cure for HIV-1.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Marc-André Langlois for welcoming me as a student into his laboratory and allowing me the opportunity to perform such rewarding research. His brilliance, availability for advice, and willingness to help have been gratefully appreciated during my many years under his supervision.

I would like to thank my Thesis Advisory Committee members, Dr. Martin Pelchat and Dr. Kathryn Wright, for their helpful discussions and questions during my TAC meetings, that brought more perspective to my project.

I would like to thank André Bergeron for his guidance in working in the CL3 lab and accompanying me to emergency when I had an accident in my lab.

I would like to thank the present and past members of my lab, Tyler Renner for his friendship and whose intelligence I admire so much, Dr. Maria Rosales Gerpe for her mentorship during my undergraduate Honours project which began my journey as a scientist, Dr. Kasandra Bélanger for being a temple of knowledge and experience whose brain I picked often, Tara Read for her kindness and readiness to help, Dr. Vera Tang for her expertise in flow cytometry, Dr. Joanne McBane for her contribution to the bioinformatics portion of my project, Richard Kang for his friendship and insightful life talks, Nebeyu Teshome and Liam Châtel for being awesome lab mates and always putting a smile on my face, Laura Goodwin and Mark Campbell for their companionship, Anna Frietzche for her friendliness, and Matthew Greig for his dedication in continuing the legacy of my project.

I would also like to thank my mom, Heber, and Timothy for their constant support during my studies and for never hesitating to drive me to the lab at odd hours to complete my experiments – God knows they had to put up with a lot! And finally, my Grandma who taught me the most important lessons in my life and who I dedicate this thesis to.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

A: Adenine
AIDS: Acquired Immune Deficiency Syndrome
AMP: Ampicillin
AP-1: Activator Protein 1
APOBEC: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like
APOBEC2, A2: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 2
APOBEC3, A3: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3
APOBEC3A, A3A: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3A
APOBEC3B, A3B: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3B
APOBEC3C, A3C: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3C
APOBEC3D, A3D: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3D
APOBEC3F, A3F: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3F
APOBEC3G, A3G: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3G
APOBEC3H, A3H: Apolipoprotein B mRNA-editing Enzyme Catalytic polypeptide-like 3H
ART: Antiretroviral Therapy
ATF/CREB: Activating Transcription Factor/cAMP Response Element Binding
AU: Arbitrary Units

bNAb: Broadly Neutralizing Antibody
bp: Base Pair

C: Cytosine
CA: Capsid
Ca: Calcium
CCR5: C-C Chemokine Receptor type 5
CD4: Cluster of Differentiation 4
CD8: Cluster of Differentiation 8
Cdk9: Cyclin-dependent kinase 9
cDNA: Complementary DNA
CDS: Coding Sequence
ChIP: Chromatin Immunoprecipitation
CMV: Cytomegalovirus
CMV-p: Cytomegalovirus Promoter
CNS: Central Nervous System
CO_2: Carbon Dioxide
COUP: Chicken Ovalbumin Upstream Promoter
cPPT: Central Polypurine Tract
CRF: Circulating Recombinant Form
CTD: C-Terminal Domain
CTF/NFI: CCAAT box-binding Transcription Factor/Nuclear Factor I
CXCR4: C-X-C Chemokine Receptor type 4
Cys: Cysteine

D: Aspartic Acid
DMEM: Dulbecco's Modified Eagle's Medium

DMSO: Dimethyl Sulfoxide
DNA: Deoxyribonucleic Acid
dNTP: Deoxynucleoside Triphosphate
dsDNA: Double-stranded DNA
DSE-1: Downstream Element 1

ECL: Enhanced Chemiluminescence
E. coli: *Escherichia coli*
EDTA: Ethylenediaminetetraacetic Acid
eGFP: Enhanced Green Fluorescent Protein
Env: Envelope
Ets-1: v-Ets erythroblastosis virus E26 oncogene homologue 1

FACS: Fluorescence-Activated Cell Sorting
FBS: Fetal Bovine Serum
FWD: Forward

G: Guanine
g: Gravity
G418: Geneticin
Gag: Group-specific Antigen
GALT: Gut-Associated Lymphoid Tissue
gDNA: Genomic DNA
GFP: Green Fluorescent Protein
Glu: Glutamic Acid
gp: Glycoprotein
gRNA: Genomic RNA

H: Histidine
HAART: Highly Active Antiretroviral Therapy
HDAC: Histone Deacetylase
HEK: Human Embryonic Kidney
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His: Histidine
HIV: Human Immunodeficiency Virus
HIV-1: Human Immunodeficiency Virus Type 1
HIV-2: Human Immunodeficiency Virus Type 2
HRM: High Resolution Melt
HRP: Horse Radish Peroxidase
HTLV: Human T-Cell Leukemia Virus

IN: Integrase
IO: Ionomycin
IPTG: Isopropyl β-D-1-thiogalactopyranoside
IST: Inducer of Small Transcripts

KAN: Kanamycin

kb: Kilobase
LB: Luria Bertani
LBP-1: Leader Binding Protein 1
LPV: Latency-Prone Virus
LRA: Latency-Reversing Agent
LTR: Long Terminal Repeat

M: Molar
MA: Matrix
mA: Milliampere
MFI: Mean Fluorescence Intensity
mg: Milligram
Mg: Magnesium
mL: Millilitre
mM: Millimolar
mRNA: Messenger RNA

n: Number
Na: Sodium
NaCl: Sodium Chloride
NC: Nucleocapsid
Nef: Negative Regulatory Factor
NFAT: Nuclear Factor of Activated T cells
NF- κB: Nuclear Factor kappa-light-chain-enhancer of activated B cells
ng: Nanogram
nM: Nanomolar
No.: Number
nt: Nucleotide
NTD: N-Terminal Domain

PBS: Primer Binding Site
PBS: Phosphate Buffered Saline
PBS-T: Phosphate Buffered Saline + Tween
PCR: Polymerase Chain Reaction
PFA: Paraformaldehyde
PIC: Pre-Integration Complex
PKC: Protein Kinase C
PMA: Phorbol Myristate Acetate
Pol: Polymerase
PPT: Polypurine Tract
PR: Protease
Pro: Proline
P-TEFb: Positive Transcription Elongation Factor
PVDF: Polyvinylidene Fluoride

qPCR: Quantitative Real-time PCR

R: Repeat Region
Rev: Regulator of Expression of Virion proteins
REV: Reverse
RF: Restriction Factor
RIPA: Radioimmunoprecipitation Assay buffer
RNA: Ribonucleic Acid
RPM: Revolutions Per Minute
RPMI: Roswell Park Memorial Institute
RRE: Rev Response Element
RT: Reverse Transcriptase
RTC: Reverse Transcription Complex

SAMHD1: Sterile Alpha Motif and Histidine/Aspartic acid domain-containing protein 1
SD: Standard Deviation
SDM: Site-Directed Mutagenesis
SDS: Sodium Dodecyl Sulphate
SDS-PAGE: Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SEM: Standard Error of the Mean
SIV: Simian Immunodeficiency Virus
Sp1: Specificity Protein 1
STI: Sexually-Transmitted Illness
ssDNA: Single-stranded DNA
SSR: Start Site Region
ssRNA: Single-stranded RNA
SU: Surface

T: Thymine
TAR: Transactivation Response
Tat: Transactivator of transcription
TF: Transcription Factor
TFBS: Transcription Factor Binding Site
TM: Transmembrane
Trim5α: Tripartite motif 5 alpha
tRNA: Transfer RNA

U: Uracil
U: Unit
U3: Unique 3’ Region
U5: Unique 5’ Region
UBP-1/LBP-1: Untranslated Binding Protein 1/Leader Binding Protein 1
UBP-2: Untranslated Binding Protein 2
USF: Upstream Stimulatory Factor
UV: Ultraviolet

V: Volts
Vif: Viral Infectivity Factor
Vpr: Viral Protein r

Vpu: Viral Protein u
VSV-g: Vesicular Stomatitis Virus glycoprotein
v/v: Volume per Volume

wt: Wildtype
w/v: Weight per Volume

X-gal: 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

°C: Degrees Celsius
µg: Microgram
µL: Microlitre
µM: Micromolar
Ψ: Psi Packaging Signal
+ssRNA: Positive-sense, single-stranded RNA
-sssDNA: Minus-strand strong stop DNA
+sssDNA: Plus-strand strong stop DNA

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

1.1 Retroviruses

Retroviruses are a unique class of RNA viruses in that they form a DNA replication intermediate (called a provirus) that becomes integrated into the genetic material of the host in which they infect. Prior to their discovery, it was believed that genetic information flowed from DNA to RNA to protein, known as the Central Dogma of molecular biology. Retroviruses challenge this idea as the information in genes flows from RNA to DNA back to RNA to protein. This is made possible due to a virally-encoded RNA-dependent DNA polymerase called reverse transcriptase (RT), which allows the virus to generate this DNA intermediate from its RNA genome (1). This DNA intermediate is then stably integrated into the host genome thereby permanently affecting the host's genetic composition. Currently, there are only two known human retroviruses that cause disease in humans: Human T-Cell Leukemia Virus (HTLV) and Human Immunodeficiency Virus (HIV) (2, 3), the latter being the virus of study in this thesis.

1.1.1 HIV-1

A new disease known as Acquired Immune Deficiency Syndrome (AIDS) was first clinically observed in 1981. Patients suffering from this syndrome were characterized by having an impaired immunity that allowed opportunistic infections to cause morbidity. The cause of the disease was unknown at the time of observation until two competing research groups, led by Robert Gallo, and Françoise Barré-Sinoussi and Luc Montagnier, independently isolated HIV as the causative agent in 1983 (4). HIV exists as two types: HIV type 1 (HIV-1) and HIV type 2 (HIV-2). Both virus types are thought to have developed from a zoonotic transmission of a recombinant Simian Immunodeficiency Virus (SIV) from non-human primates to humans in Central and Western Africa (5, 6). HIV-1 is the more virulent and

transmittable of the two types and is the one we associate most with the HIV/AIDS pandemic. It is the result of a zoonotic transmission of SIV from chimpanzees, whereas HIV-2, the less virulent and less transmittable of the two, is the result of a zoonotic transmission of SIV from sooty mangabeys (5, 6). As of 2017, there is an estimated 36.9 million people worldwide currently living with HIV/AIDS (7). Since its identification, it has taken a death toll of over 35 million people worldwide (7).

1.1.1.1 Genome and Virion

HIV-1 is a lentivirus characterized for its slow and chronic infection. It is an enveloped retrovirus with a positive-sense, single-stranded RNA (+ssRNA) genome. Its genome size is roughly 9.7 kb and encodes for nine genes which produce a total of fifteen proteins (Figure 1.1). Every retrovirus codes for essential Gag, Pol, and Env polyproteins. The Gag polyprotein is further cleaved into the nucleocapsid (NC), capsid (CA), and matrix (MA) structural proteins (8). The Pol polyprotein gives rise to the reverse transcriptase (RT), integrase (IN), protease (PR), and RNAse H enzymes (8). The Env polyprotein is processed and cleaved into two envelope glycoproteins, gp41, a transmembrane (TM) subunit, and gp120, a surface (SU) subunit (8, 9). HIV-1 is a complex retrovirus and also codes for six accessory proteins: the regulatory proteins, Transactivator of transcription (Tat) and Regulator of expression of virion proteins (Rev); and the accessory proteins, Viral infectivity factor (Vif), Viral protein r (Vpr), Viral protein u (Vpu), and Negative regulatory factor (Nef). The entire proviral genome is flanked by two 634 bp Long Terminal Repeats (LTRs), which serve as the promoter at the 5' end of the genome as well as polyadenylation and transcription termination at the 3' end of the genome (10). Both LTRs are organized in the same architecture, composed of U3, R, and U5 regions. The U3 region houses the majority of the transcription factor binding sites (TFBS) and contains the promoter, enhancer, and other regulatory sequences. The R region marks the

start of the viral transcript at the 5’ end and contains the transactivation response (TAR) element. At the 3’ end, the R region is also the site of polyadenylation and transcription termination.

The HIV-1 genetic material itself is present as two copies of single-stranded RNA contained within a conical-shaped core formed from the CA protein (Figure 1.2). The two viral RNA copies, along with viral enzymes RT, IN, and PR, are held together by their association with the NC protein. The viral core is surrounded by protective MA protein underneath a lipid bilayer envelope that the virus snatches from the host cell membrane. The envelope glycoproteins, present as heterotrimers, are then embedded and protrude out of the virus, as well as any host proteins that were picked up upon viral egression (8, 11–13).

Figure 1.1. HIV-1 Genome Organization. The HIV-1 RNA genome begins from the R region of the 5' LTR and ends after the R region of the 3' LTR. Upon reverse transcription, a double-stranded DNA intermediate is formed with duplicates of the LTRs containing U3, R, and U5 regions at both ends of the viral DNA genome. Depicted are the primer binding site (PBS), Psi packaging signal (Ψ), central polypurine tract (cPPT), Rev response element (RRE), 3' polypurine tract (3'PPT), as well as the organization of the various genes and proteins encoded by the virus. Reproduced from Reference (14) (Open Access).

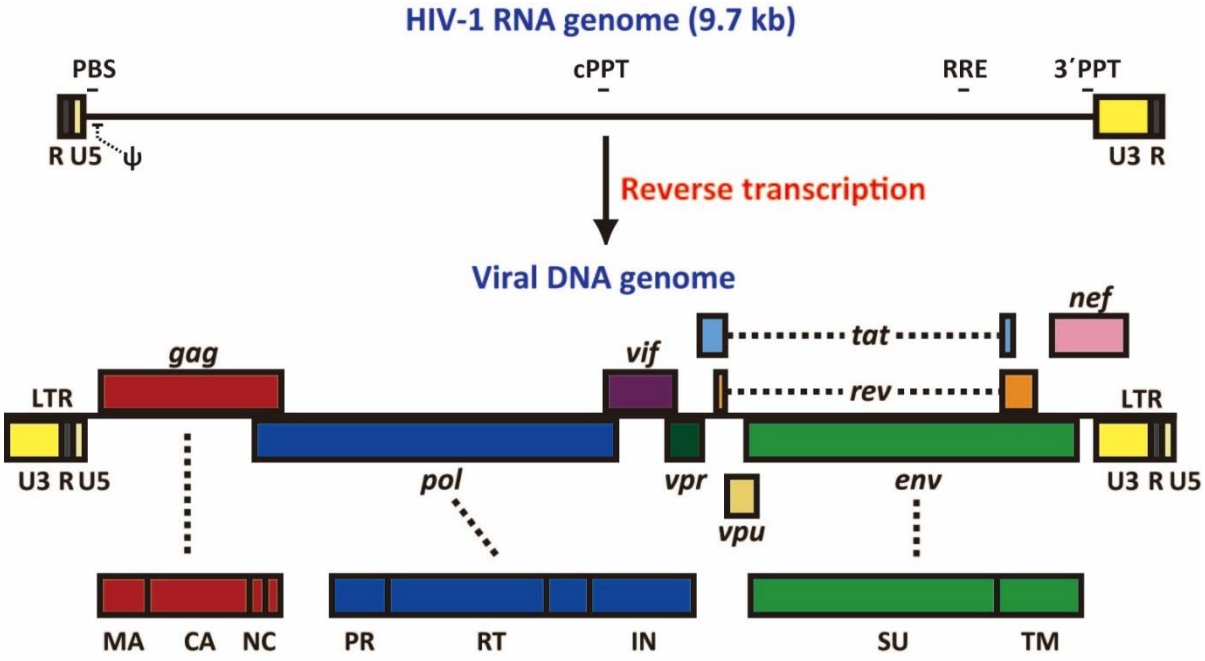
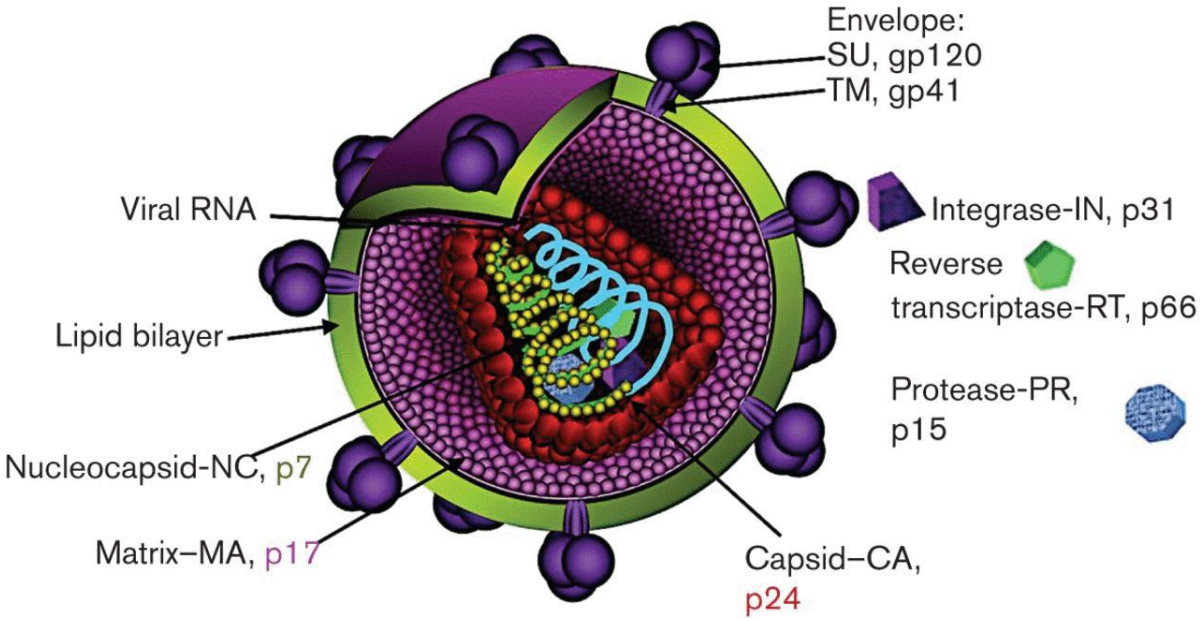

HIV-1 RNA genome (9.7 kb)
PBS
cPPT
RRE
3´PPT
R U5
ψ
U3 R
Reverse transcription
Viral DNA genome
nef
tat
gag
vif
rev
LTR
LTR
U3 R U5
pol
vpr
env
U3 R U5
vpu
MA
CA
NC
PR
RT
IN
SU
TM

Figure 1.2. HIV-1 Virion Structure. Schematic representation of a mature HIV-1 virion. Reproduced from Reference (15) (Open Access).

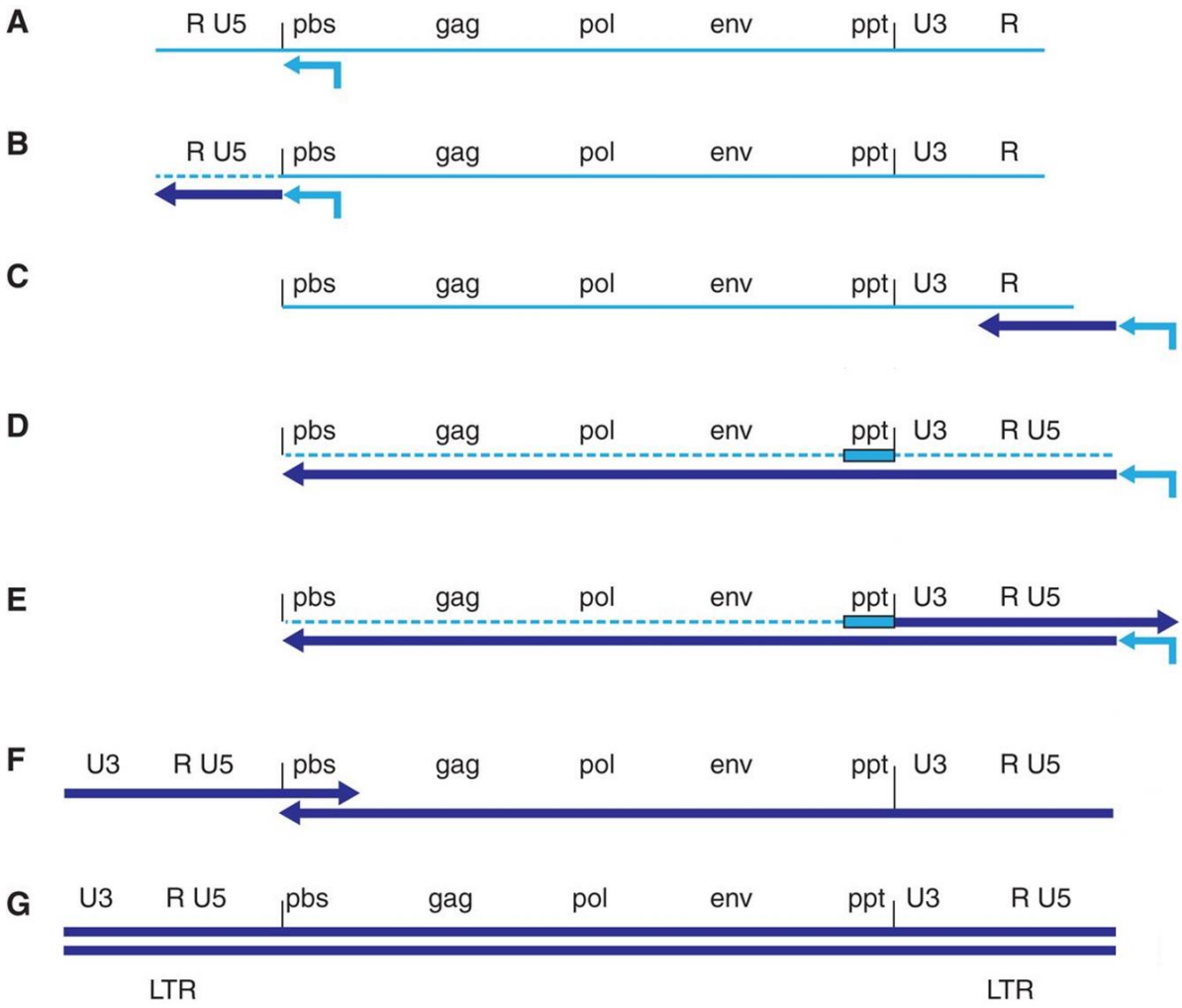

Envelope:
SU, gp120
TM, gp41
Viral RNA
Integrase-IN, p31
Reverse
transcriptase-RT, p66
Lipid bilayer
Protease-PR,
p15
Nucleocapsid-NC, p7
Matrix–MA, p17
Capsid–CA,
p24

1.1.1.2 Replication Cycle

The infection of HIV-1 begins when the virus finds a suitable target host cell. HIV-1 primarily infects cells of the immune system such as T lymphocytes, macrophages, dendritic cells, and microglial cells (16–20). These cells express CD4 receptors on their surface that the viral gp120 SU glycoprotein recognizes and interacts with (21). Upon interaction, a conformational change occurs, allowing for the binding to a chemokine co-receptor, either CXCR4 or CCR5 (these also dictate the tropism of the virus) (22–26). Once the receptors are bound, the gp41 TM glycoprotein is then inserted into the host cell membrane, leading to fusion of the two membranes and viral entry into the host cell (21, 27–29). The viral capsid remains intact and reverse transcription of the viral genomic RNA takes place inside this core (forming a reverse transcription complex (RTC)) (30–33). See Figure 1.3 for a depiction of the reverse transcription process. The generation of the minus-strand complementary DNA (cDNA) from the viral genomic RNA is primed by a host $tRNA_3^{Lys}$ primer on the primer binding site (PBS) located just downstream of the 5' LTR. Reverse transcriptase (RT) works in a 3' to 5' orientation and synthesis of the minus-strand temporarily stops at the 5' end of the viral RNA genome. This short fragment is known as the minus-strand strong stop DNA (-sssDNA) and requires a template switch, known as the minus-strand transfer, to complete its synthesis. The strand is transferred to the 3' end of the same viral RNA strand or to the second copy of the genome, where the newly synthesized R region of the -sssDNA binds to its complementary sequence at this end. Minus-strand synthesis continues from the R region to the PBS and at the same time, the RNAse H activity of the RT digests the RNA in the RNA:DNA hybrids created. However, digestion is incomplete as some fragments remain intact, namely the polypurine tracts (PPTs). HIV-1 has two PPTs, a central PPT (cPPT) and a 3' PPT. The latter is essential and used as a primer during reverse transcription for plus-strand

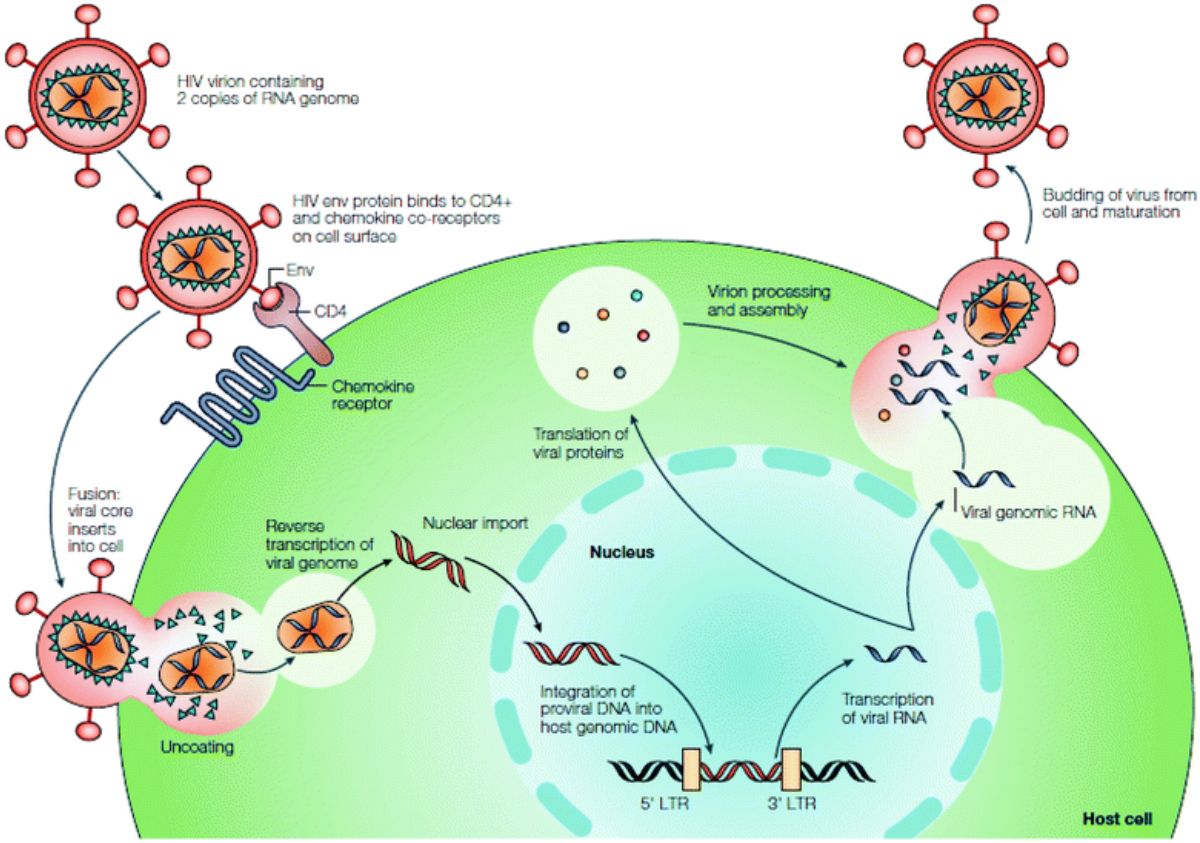
synthesis. Plus-strand synthesis proceeds from the 3’ PPT to the 5’ terminus of the minus-strand. The small fragment produced is known as the plus-strand strong stop DNA (+sssDNA) and similarly, a strand transfer must occur to continue its synthesis. This plus-strand transfer allows the PBS of the +sssDNA to anneal to the PBS of the minus-strand. Both DNA strands then use each other as templates to complete their own sequence, resulting in two identical strands of viral DNA flanked by complete LTRs at each end. This is known as the provirus and is now ready for nuclear import. Once reverse transcription is complete, the viral core containing the provirus is trafficked to the nuclear pore where it is uncoated and the provirus becomes part of a pre-integration complex (PIC) (32). The PIC enters the nucleus and with the aid of IN and other cellular proteins, the provirus is then integrated into the host genome (31, 34–36).

Figure 1.3. HIV-1 Reverse Transcription. (A) Minus-strand synthesis begins by annealing of a host tRNA primer to the primer binding site (pbs) of the viral RNA (light blue line). (B) Minus-strand synthesis progresses until the 5' end of the viral genome (cDNA shown as dark blue line). This short fragment is known as the minus-stand strong stop DNA (-sssDNA). The RNA in the RNA:DNA duplex is digested by the RNAse H activity of RT. (C) Minus-strand transfer anneals the newly-formed R region of the nascent minus-strand complementary to the R region at the 3' end of the viral genome. (D) Minus-strand synthesis continues until the end of the remaining viral genome. RNAse H activity digests the RNA, leaving behind the polypurine tract (ppt). (E) The ppt is used as the primer for plus-strand synthesis. Plus-strand synthesis progresses until it reaches the sequence of the tRNA primer. This short fragment is known as the plus-strand strong stop DNA (+sssDNA). (F) Plus-strand transfer anneals the newly-formed pbs of the nascent plus-strand complementary to the pbs of the minus-strand. The ppt and tRNA primer are digested by RNAse H. (G) Synthesis of the remainder of each strand proceeds until a double-stranded DNA intermediate is produced of the viral genome flanked by complete LTRs (U3, R, U5). Reproduced from Reference (37) with permission from Cold Spring Harbor Laboratory Press.


A
R U5
pbs
gag
pol
env
ppt
U3
R
B
R U5
pbs
gag
pol
env
ppt
U3
R
C
pbs
gag
pol
env
ppt
U3
R
D
pbs
gag
pol
env
ppt
U3
R U5
E
pbs
gag
pol
env
ppt
U3
R U5
F
U3
R U5
pbs
gag
pol
env
ppt
U3
R U5
G
U3
R U5
pbs
gag
pol
env
ppt
U3
R U5
LTR
LTR

Once the provirus is integrated into the host genome, it must depend on host cell machinery to express its viral genes and replicate its own viral RNA genome (31, 38). Viral transcription is driven by the 5' LTR promoter. Cellular transcription factors (TFs) are recruited to the site of transcription and initially, only short mRNA transcripts are generated (39, 40). It is not until the accessory protein Tat is produced in sufficient amount that viral transcription progresses efficiently (39, 40). Tat returns to the nucleus and binds a stem-loop structure on the viral RNA called the transactivation response (TAR) element. This leads to the recruitment of the positive transcription elongation factor (P-TEFb) complex, comprised of cyclin-dependent kinase 9 (Cdk9) and cyclin T1, which then phosphorylates the C-terminal domain (CTD) of RNA polymerase II (41). This increases the processivity of the enzyme and allows it to get over the transcription block. Tat transactivation thereby vastly improves transcription elongation and efficiency. Another accessory protein, Rev, also returns to the nucleus where it binds to the Rev response element (RRE) on newly generated transcripts. This allows for the export of unspliced and incompletely spliced transcripts out to the cytoplasm (42). Once in the cytoplasm, translation of viral proteins occurs and new virions are assembled at the cell surface. Full-length viral RNA genomes are packaged into viral particles by their association with the NC domain of Gag via the psi (Ψ) packaging signal located downstream of the 5' LTR (43). The progeny viruses then bud out of the host cell, taking the host cellular membrane with it, studded with viral envelope glycoproteins that it catches along the way (27–29). Once outside of the cell, the virus is fully matured by protease cleavage of the Gag and Pol polyproteins by the viral PR, which then generates the final conical capsid shape and viral proteins necessary for the progeny viruses to infect new cells (Figure 1.4) (44, 45).

Figure 1.4. HIV-1 Replication Cycle. Schematic representation of the steps in the HIV-1 replication cycle beginning with binding of the virus to host cell receptors and ending with budding of the virus from the host cell and virus maturation. Reproduced from Reference (46) with permission from Springer Nature.


HIV virion containing
2 copies of RNA genome
HIV env protein binds to CD4+
and chemokine co-receptors
on cell surface
Env
CD4
Chemokine
receptor
Fusion:
viral core
inserts
into cell
Uncoating
Reverse
transcription of
viral genome
Nuclear import
Nucleus
Integration of
proviral DNA into
host genomic DNA
5' LTR
3' LTR
Transcription
of viral RNA
Translation of
viral proteins
Virion processing
and assembly
Viral genomic RNA
Budding of virus from
cell and maturation
Host cell

1.1.1.3 Pathogenesis

HIV-1 infection is primarily a sexually-transmitted illness (STI). However, it can be transmitted through various forms of contact with infected bodily fluids, such as through blood transfusions, unsterile drug use, vertically through childbirth, etc (47–49). As mentioned previously, HIV-1 targets cells of the immune system decorated with the CD4 receptor. This includes $CD4^+$ T lymphocytes, macrophages, dendritic cells, and microglial cells. Co-receptor binding determines the tropism of the virus. The chemokine co-receptor CCR5 is used to mainly infect macrophages and memory $CD4^+$ T cells (denoted M-tropic, R5 viruses) during initial infection and it is thought that macrophages harbour the initial viruses and contribute to their production and expansion when $CD4^+$ T cells start to become depleted (50). Later on during infection, there is a switch in co-receptor preference and the chemokine co-receptor CXCR4 is used to mainly infect a variety of peripheral T cells (denoted T-tropic, X4 viruses) (50). This explains the more adverse impact of the infection later on during its course as T cells continue to deplete, ultimately leading to the development of AIDS.

There are two clinical phases in an HIV-1 infection, an acute phase and a chronic phase, ending with the progression to AIDS. During the acute phase, a patient may experience flu-like symptoms as the virus undergoes intense replication inside the host cells (51). A huge portion of the $CD4^+$ T cell population in the peripheral blood and gut-associated lymphoid tissue (GALT) becomes depleted during this time due to cytopathic effects from virus production, immune response through cytotoxic $CD8^+$ T cells, inflammation, etc (6, 52–56). Following this acute phase, levels of $CD4^+$ T cells appear to bounce back in the bloodstream, however, cell numbers in the GALT remain depleted throughout the infection (54–56). Viral levels peak and then decline, reaching a level that allows the virus to reside in the host for many years, spreading slowly. During the chronic phase, $CD4^+$ T cells continue to be depleted

and their regeneration declines, leading to immune dysfunction and increased susceptibility to opportunistic infections, a hallmark of AIDS (57).

In its early days following the discovery of the virus, HIV-1 infection and progression to AIDS meant certain death. However, with the advent of antiretroviral therapy (ART), the life expectancy of an HIV-1 patient is now comparable to that of the general population (58–60). Today, HIV-1-infected patients undergo a regime known as highly active antiretroviral therapy (HAART), a combination therapy consisting of at least three different antiretroviral drugs that target different stages of the viral infection cycle (45, 61, 62). Because of HIV-1's rapid evolution, owing in part to the poor fidelity of the RT enzyme (which can introduce up to 10 mutations per genome synthesized), it is not surprising that the virus has evolved drug resistance mutations (37, 62–68). A combination therapy therefore helps to circumvent this problem by reducing the chances of developing resistance because of specific mutations. HAART has the capacity to reduce plasma viral loads to below detection limits (<50 copies/mL) and keep it suppressed for the duration of treatment, thereby greatly enhancing the quality of life of infected individuals (69–71).

1.2 Latency and Persistence

Despite the effectiveness of HAART, these life-saving drugs are often very toxic to the patient in the long-term. Therefore, the current goal to improve HIV-1 therapy is to be able to take the patient off of HAART eventually. However, when HAART is ceased, there is rapid rebound in the plasma viral loads observed in the patients (72–74). HAART is unable to completely eradicate HIV-1 infection because the virus is able to establish a state of latent infection within cellular reservoirs (41). In this latent state, the replication-competent provirus has successfully integrated into the host genome; however, the virus lies dormant and very little transcription takes place (17, 41). This leads to undetectable levels of viral protein or viral

particle production, preventing the host immune system from mounting a response to successfully clear the pathogen (17). Similarly, antiretroviral therapeutic approaches, which target actively replicating virus, are also unable to eliminate the infection (17). The latent virus is then able to hide unscathed until the cell becomes activated, thereby causing HIV-1 transcription to reinitiate. The newly produced virus particles then contribute to further spread the infection and re-seed the viral reservoir. This in fact constitutes the biggest obstacle to curing HIV-1. Furthermore, the phenomenon of viral latency is not an evolutionary accident for the virus and is not merely dependent on cell cycle state (75, 76). There is increasing evidence that viral latency is a hardwired strategy embedded in the viral makeup to allow for viral persistence and enhance viral transmission (75, 76). What is also troubling is the possibility that the latent reservoir is even greater than initially imagined. One study found that even after maximum *in vitro* T cell activation, less than 1% of proviruses were induced to produce infectious virus (77). Although much of this can be attributed to defects in the genome of the non-induced proviruses, 11.7% of them were found to be replication-competent, meaning that they could have the potential to become activated *in vivo* (77). This would drastically increase the size of the latent reservoir (~60-fold) and thus complicate finding a cure for HIV-1 (77).

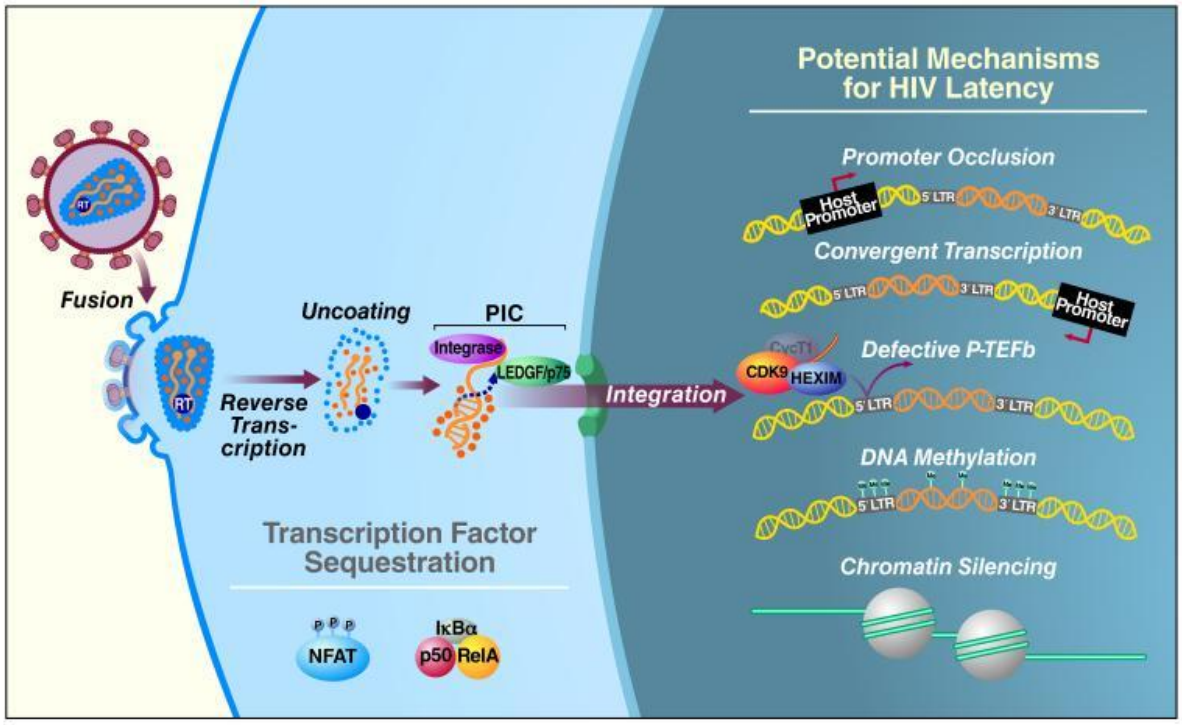
1.2.1 Current Models

HIV-1 primarily establishes a latent reservoir within resting memory $CD4^+$ T cells (41, 78–81). Latent HIV-1-infected cells represent ~1 in 1 x 10^6 cells in this population (75, 82). These cells are long-lived, allowing the virus to persist in a latent state within the host for many years. Early on during HIV-1 infection, activated $CD4^+$ T cells become infected (72, 74, 83, 84). As the infection progresses, most of the infected cells die off. However, a subset of these cells survive long enough to revert to a resting memory state (41). This state is not conducive

to viral gene expression due, in part, to the lack of key host cellular transcription factors as well as viral regulatory proteins, such as Tat (41, 78, 85). This results in decreased viral transcription and viral evasion of the host immune system. HIV-1 also sets up infection in macrophages, dendritic cells, and microglial cells and the presence of latent reservoirs in these cell populations is still being researched (17–20, 35, 79, 86, 87). Along with these cell types, latent reservoirs have also been suggested in privileged anatomical sites such as the central nervous system (CNS) and testes where drug penetration is limited (35, 88–90). The mechanisms underlying viral latency are complex and poorly understood and are most likely due to a combination of factors. Current models for the establishment and maintenance of latency include several molecular and epigenetic mechanisms (Figure 1.5) (35, 91). Chromatin silencing through epigenetic modifications such as histone deacetylation at the site of provirus integration, and DNA methylation of the provirus, are popular mechanisms proposed for HIV-1 latency as they both lead to transcriptional silencing (35, 41, 79, 91, 92). Sequestering of key transcription factors in the cytoplasm of the cell, making them unavailable in the nucleus for transcription also leads to silencing of viral gene expression (41, 79). The orientation of provirus insertion can result in two scenarios of transcriptional interference: promoter occlusion or convergent transcription (17, 35, 41, 79, 85, 93). In the first scenario, a host gene promoter is located upstream of the viral promoter, causing the host RNA polymerase II to read through the proviral promoter and displace bound transcription factors. In the second scenario, the host gene and the provirus are in opposite orientation, causing the RNA polymerase II complexes to collide, leading to early transcription termination for both promoters or whichever is weakest. As mentioned previously, Tat transactivation is important for the generation of full-length viral transcripts. Cyclin T1, one of the proteins that make up the P-TEFb complex that is recruited by Tat, is present in low amounts in resting $CD4^+$ T cells

(41). This results in a defective P-TEFb complex, leading to lowered Tat transactivation and the promotion of HIV-1 latency (35, 41, 79, 92).

Figure 1.5. Mechanisms of HIV-1 Latency Establishment. Some of the major mechanisms of HIV-1 latency present in $CD4^+$ T cells. See text for details. Reproduced from Reference (41) with permission from Elsevier.


Potential Mechanisms for HIV Latency
Promoter Occlusion
Host Promoter
5' LTR
3' LTR
Convergent Transcription
Defective P-TEFb
CycT1
CDK9
HEXIM
DNA Methylation
Chromatin Silencing
Fusion
RT
Reverse Trans-cription
Uncoating
PIC
Integrase
LEDGF/p75
Integration
Transcription Factor Sequestration
NFAT
IκBα
p50
RelA

1.2.2 Latency Reversal

Due to the importance and burden viral latency places on developing a cure for HIV-1, many strategies have been employed in an effort to reverse the latent state of the virus. A popular strategy known as the "shock-and-kill" method uses latency reversing agents (LRAs) to stimulate the virus out of its latent state, and then relies on the host immune system or antiretroviral therapy to clear the activated viruses (17, 79, 94–96). Common LRAs include protein kinase C (PKC) agonists, which stimulate the NF-κB signalling pathways to increase transcription and gene expression (97); histone deacetylase (HDAC) inhibitors, which block the deacetylation of histones thereby opening up the chromatin for transcription (94); DNA methyltransferase inhibitors, which have a similar function of making the proviral DNA accessible for transcription (94); as well as other compounds that indirectly activate the NF-κB pathway, leading to T cell activation. Broadly neutralizing antibodies (bNAbs) have also been investigated for their use alongside LRAs to purge the latent reservoir (98). Although these LRAs and their combinations have had varied success in studies done *in vitro*, their ability to cure latently infected cells *in vivo* in clinical trials has so far proven ineffective (94, 99). Therefore, other strategies of latency reversal are also being explored. These include gene therapy/gene editing and stem cell transplantation to delete the virus from the cells or to make them more resistant to HIV-1 infection (92, 100, 101). However, these approaches are far from being easily accessible treatments and their safety and efficacy still need to be studied further.

1.3 Intrinsic Immunity

HIV-1, although slow in its method of causing chronic, persistent infections, is very fast to infect. It only takes a few hours for the virus to get into the cell and stably integrate its genome into the host chromosome. Innate and adaptive immune responses, which take from a few days to a few weeks, respectively, to kick in, are thus useless in the initial protection

against this virus. Fortunately, we have intrinsic immune defences that come into play when encountering these retroviral pathogens. We possess in our cells, host-encoded retroviral restriction factors (RFs), that are constitutively expressed and act immediately upon retroviral infection. They represent our first line of defence against retroviruses until the innate and adaptive immune responses become activated later on. The most effective restriction factors against retroviruses are Trim5α, SAMHD1, Tetherin, and APOBEC3.

1.3.1 Restriction Factors

Trim5α (tripartite motif 5 alpha) is a cellular protein that acts by directly binding to the capsid protein of the retrovirus and orchestrating its premature uncoating once the virus has entered the host cell (102–106). This leads to the disruption of reverse transcription as well as nuclear import (105). It provides cross-species protection against retroviruses but is poorly active against retroviruses of the same host species (107, 108). SAMHD1 (sterile alpha motif and histidine/aspartic acid domain-containing protein 1) acts by depleting intracellular pools of deoxynucleoside triphosphates (dNTPs) in myeloid lineage cells (105, 106, 109). HIV-1 has the special ability to replicate in non-dividing cells and cells with limited concentrations of dNTPs, such as resting memory $CD4^+$ T cells, macrophages, and dendritic cells (110–113). SAMHD1 further depletes the already-low pools of dNTPs in these cells, making them unavailable for reverse transcription (114). Tetherin is a type II transmembrane protein that acts by anchoring the progeny viruses to the surface of the cell and preventing their release (105, 106, 115). The viruses are then internalized by endocytosis and degraded by lysosomes (116, 117).

1.3.2 APOBEC3

APOBEC3 (apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3, A3) proteins are a family of cytidine deaminases that mutate deoxycytidines into deoxyuridines on

single-stranded DNA (118–120). Humans encode for seven APOBEC3 genes located in tandem on chromosome 22, denoted APOBEC3A (A3A), APOBEC3B (A3B), APOBEC3C (A3C), APOBEC3D (A3D), APOBEC3F (A3F), APOBEC3G (A3G), and APOBEC3H (A3H), with only the latter four proteins functioning against HIV-1 *in vivo* (119, 121–124). The prior three are not present at physiological levels or do not encapsidate into viral particles or restrict retroviral infection in relevant host cells (125). The APOBEC3 proteins act on single-stranded DNA in a 3' to 5' processive manner (126). A3F and A3G are the most potent of the four A3s against HIV-1 infection and have differing deamination dinucleotide contexts. A3G prefers to deaminate cytosines (C) preceded by another cytosine, in a 5'CC deamination target motif, whereas A3F and all other A3s prefer to deaminate cytosines preceded by a thymine (T), in a 5'TC deamination target motif (121, 126–128).

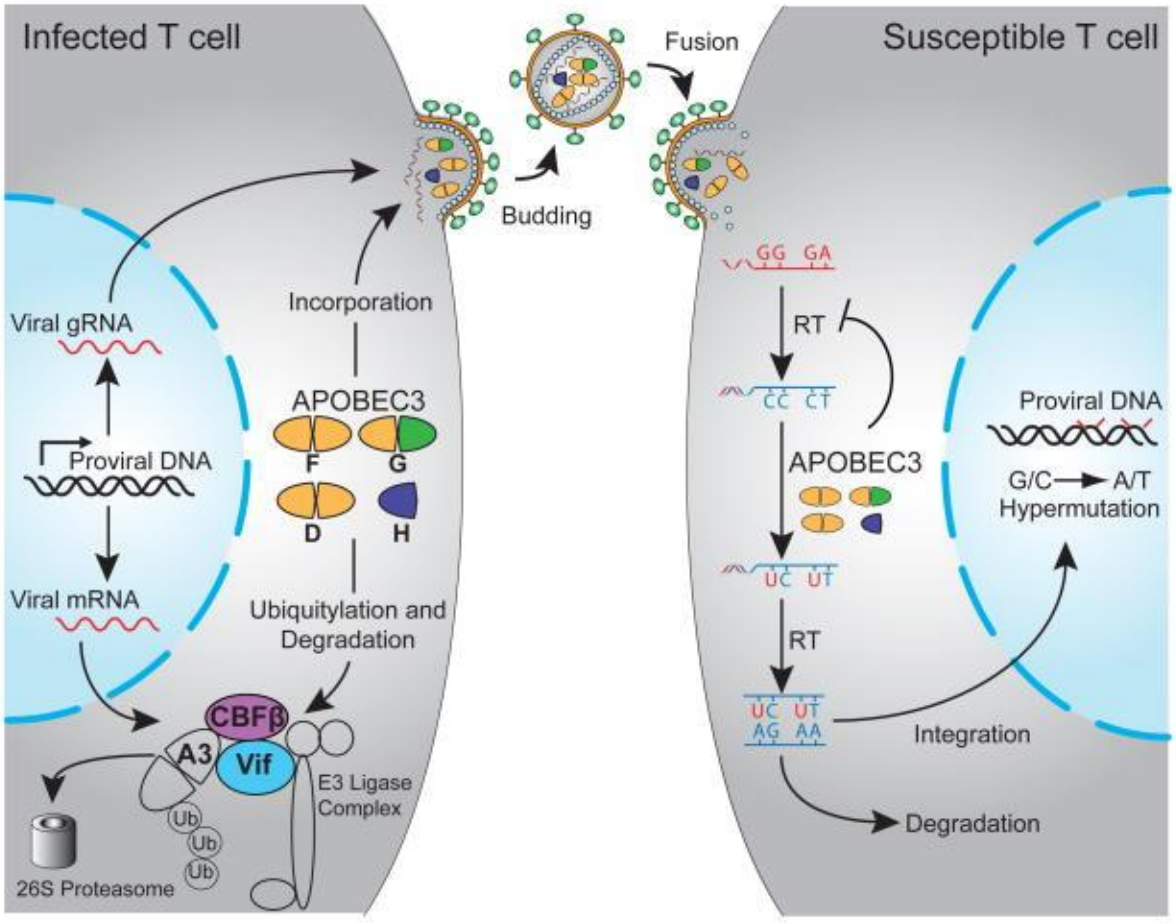
The discovery of the A3 proteins came with the observation that certain cell types, termed non-permissive cells, were not susceptible to HIV-1 infection in the absence of its accessory protein Vif due to a factor present in the cells. Other cell lines, termed permissive cells, were able to be infected by this same virus and allowed for the spread of HIV-1 infection even in the absence of Vif. A study using cDNAs from a permissive and non-permissive cell line as probes in a northern blot analysis of RNAs isolated from various permissive and non-permissive cells was performed by Sheehy and colleagues to elucidate the factor responsible (129). The experiments enriched for a particular cDNA that was expressed in all of the non-permissive cells but was absent or minimally-expressed in the permissive cells (129). This protein was identified as A3G (at the time called CEM15) and soon after, other members of the family were discovered based on homology (130, 131). The A3 proteins are characterized by a conserved zinc-binding motif (Cis/His)-Xaa-Glu-$Xaa_{23\sim28}$-Pro-Cys-$Xaa_{2\sim4}$-Cys, where Xaa can be any amino acid, in their active sites (119, 130, 131). A3A, A3C, and A3H have

one zinc-coordinating domain, while A3B, A3D, A3F, and A3G have two (131). In A3G, the C-terminal domain (CTD) is responsible for its deamination activity, while the N-terminal domain (NTD) is responsible for non-catalytic activities such as RNA binding, viral encapsidation, and Vif interaction (132, 133). For A3F, the CTD is responsible for all of these functions (132, 134, 135).

1.3.2.1 Deaminase-Dependent Restriction

In Vif-deficient HIV-1 infection of producer cells, A3 proteins in the cytoplasm are packaged into budding progeny viruses along with the viral RNA. Upon reaching a target cell, the A3s begin to deaminate deoxycytidines into deoxyuridines on the single-stranded nascent minus cDNA during reverse transcription (136–140). During synthesis of the complimentary plus-strand, these uridines (U) template the incorporation of adenines (A) in the place of what would have been guanines (G), and for this reason, the A3 proteins are described as causing G-to-A hypermutation of the proviral DNA (Figure 1.6). This hypermutation ultimately leads to degradation of the viral genome, or if the provirus is able to integrate, leads to lowered fitness of the virus.

Figure 1.6. APOBEC3 Deamination of HIV-1. APOBEC3 in the cytoplasm of an infected cell gets packaged along with the viral RNA in budding progeny virions. Upon viral entry into a susceptible target cell, the encapsidated A3 proteins deaminate cytosines into uracils on the nascent minus-strand cDNA during reverse transcription. The uracils template the incorporation of adenines in the complementary plus-strand during synthesis and leads to G-to-A hypermutation of the proviral DNA. The hypermutated proviral DNA is then degraded or integrated into the host genome. When the viral protein Vif is present in the infected producer cell, it antagonizes A3 by binding to it in the cytoplasm and tagging it for proteasomal degradation. Reproduced from Reference (124) with permission from Elsevier.


Infected T cell
Fusion
Susceptible T cell
Budding
Viral gRNA
Incorporation
APOBEC3
F
G
D
H
Proviral DNA
Viral mRNA
Ubiquitylation and
Degradation
CBFβ
A3
Vif
E3 Ligase
Complex
Ub
Ub
Ub
26S Proteasome
GG GA
RT
CC CT
APOBEC3
UC UT
RT
UC UT
AG AA
Integration
Degradation
Proviral DNA
G/C → A/T
Hypermutation

1.3.2.2 HIV Antagonism of Restriction Factors

In HIV-1 infections where Vif is present, this viral accessory protein antagonizes A3 by preventing its encapsidation into progeny viruses (106, 136, 141). It does so by binding to A3 in the cytoplasm and with the aid of the E3 ubiquitin ligase complex, ubiquitylates it for proteasomal degradation through the 26S proteasome (142, 143). Similarly, HIV's other accessory proteins, namely Vpu and Vpx (in HIV-2), also have roles antagonizing specific restriction factors. Vpu in HIV-1 antagonizes tetherin by binding to the restriction factor and downregulating its expression at the cell surface (106, 144). This prevents it from anchoring progeny viruses to the cell membrane, thereby allowing efficient release of the viral particles. In HIV-2, Vpx is encoded instead of Vpu. It antagonizes SAMHD1 by inducing its proteasomal degradation, thereby freeing up dNTPs for use in reverse transcription (106, 145).

1.3.2.3 Deaminase-Independent Restriction

Although the A3s mainly restrict retroviral infection through deamination-dependent mechanisms, they have also been shown to restrict retroviral infection in the absence of catalysis *in vitro* (146–151). The mechanisms of deamination-independent restriction have not been fully elucidated but it has been proposed that the A3s can inhibit reverse transcription of the virus by interfering with primer annealing and reverse transcription initiation, as well as through steric hindrance by directly binding to the viral genomic RNA or the RT (124, 152–157). They can also inhibit integration of the virus by interacting with the IN protein during the integration process (158).

1.3.2.4 A Double-Edged Sword

As mentioned previously, the A3 proteins are antagonized by the viral protein Vif. However, Vif is a protein that accumulates late in the infection due to the requirement of Rev for its expression (125). As levels of Vif increase over time, the number of cytosolic A3

proteins available to be packaged into progeny viruses decreases, resulting in a short time frame of opportunity for sublethal mutagenesis of the virus (149, 159–162). It has been shown that sublethal mutation levels can sometimes benefit the virus and increase its overall fitness (163–168). This can be achieved by introducing resistance mutations to drugs, by modulating their viral epitopes to escape immune clearance, and by increasing genetic diversity (65, 162, 163, 165, 166, 169–171). In addition, since we specifically look at mutations caused by A3G and A3F in this thesis, it is interesting to note that deamination by A3G is often more deleterious, given its preferred deamination target motif, as it has the capacity to introduce premature stop codons at high frequency, whereas deamination by A3F introduces premature stop codons less frequently and has the possibility of generating sublethal mutations pivotal to the genetic evolution of the virus (160, 164, 167, 168, 172–174).

2.0 RATIONALE, HYPOTHESIS, AND OBJECTIVES

2.1 Rationale

Given that APOBEC3 proteins can mutate HIV-1 and it has been shown in the literature that in some instances sublethal mutation levels are compatible with viral infectivity, this data supports the possibility that mutations in regulatory regions of HIV-1 may also modulate virus expression, and consequently enable a state of proviral latency.

2.2 Hypothesis

We hypothesize that low levels of APOBEC3-induced mutations in the HIV-1 LTR promoter generate viruses with compromised transcriptional activity that exhibit a latency phenotype.

2.3 Objectives

1. Create a library of HIV-1 LTR clones mutated by APOBEC3F and APOBEC3G.
2. Examine how these mutations affect promoter expression and transactivation by Tat.
3. Investigate how A3 mutations in the LTR promoter of HIV-1 impact proviral expression and responsiveness to induction.

3.0 MATERIALS AND METHODS

3.1 Cell Culture

Human Embryonic Kidney epithelial (HEK 293T and HEK 293) cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM)/High Glucose medium (Multicell, Wisent Inc., St-Bruno, Quebec, Canada) supplemented with 10% fetal bovine serum (FBS) (CellGro, Mediatech, Inc., Manassas, Virginia, USA), 100 U/mL penicillin and 100 µg/mL streptomycin (penicillin/streptomycin) (Multicell). Human $CD4^+$ T lymphocyte (CEM-SS) cells were maintained in RPMI 1640 medium (Multicell) supplemented with 10% FBS and 1% penicillin/streptomycin. Human osteosarcoma (GHOST $CD4^+CXCR4^+CCR5^+$) cells were maintained in DMEM/High Glucose medium supplemented with 10% FBS, 1% penicillin/streptomycin, 500 µg/mL G418, 100 µg/mL hygromycin, and 1 µg/mL puromycin. Stably-transfected HEK 293 cells were kept under selection with 750 µg/mL G418 in addition to normal culture medium. All cell lines were propagated in T-75 culture flasks (Corning Inc., Tewksbury, Massachusetts, USA) at 37 °C in a 5% CO_2 incubator (Forma Series II Water Jacket CO_2 Incubator) (Thermo Scientific Inc., Burlington, Ontario, Canada).

3.2 Plasmids

The viral vector encoding single-round HIV-1 (pNL4-3-deltaE-EGFP) was obtained through the NIH AIDS Reagent Program (Catalog #11100). The *Env* gene of this vector has been disrupted with the insertion of an enhanced green fluorescent protein (eGFP) reporter gene and rendered non-functional. Our lab has further manipulated this vector by inactivating the *Vif* gene via two stop codons introduced by site-directed mutagenesis (SDM) (175) and herein referred to as pNL4-3-deltaE-EGFPΔVif. The viral vectors encoding replicative HIV-1 (pNL4.3-GFP_IRES_Nef WT and pNL4.3-ADA-GFP_IRES_Nef WT) have been graciously gifted by Dr. Éric Cohen (Institut de Recherches Cliniques de Montréal) and

express the green fluorescent protein (GFP) reporter gene in place of the *Nef* gene (176, 177). The latter vector is R5-tropic through the insertion of ADA Env sequence into the previous vector (178, 179). The Vesicular Stomatitis Virus envelope glycoprotein (VSV-g) expression plasmid used to pseudotype single-round HIV-1 infections was obtained through Addgene (Cambridge, Massachusetts, USA). The expression plasmid encoding the viral protein Tat was a gift from Dr. Paul MacPherson of the Ottawa Hospital Research Institute. An HcRed-expressing Tat plasmid was generated by inserting the Tat coding sequence (CDS) into an HcRed expression plasmid (Clontech, Takara Bio USA, Inc., Mountain View, California, USA) using the primers HcRedC1-Tat SacII FWD and HcRedC1-Tat BamHI REV (Table A1 in Appendix). Successful insertion was confirmed through Sanger sequencing using the primers Seq Tat in HcRedC1 FWD and Seq Tat in HcRedC1 REV (Table A1). The human APOBEC2, APOBEC3G, and APOBEC3F expression plasmids have been described previously (175). FLAG-tagged APOBEC proteins were used for transfection and infection assays, while eGFP-tagged APOBEC proteins were used to generate APOBEC stably-transfected HEK 293 cell lines.

3.3 Generation of Mutated HIV-1 LTR Libraries

3.3.1 Transfection and Infection

HEK 293T cells were used to produce the infectious VSVg-pseudotyped HIV-1 viral particles with packaged APOBEC3 proteins by co-transfection. 3.0×10^5 cells were seeded per well of a 6-well plate in 2 mL of DMEM and incubated for approximately 24 hours until they reached 80% confluency. Prior to transfection, the media was replaced with 3 mL of fresh DMEM. Co-transfections were performed with 500 ng of pNL4-3-deltaE-EGFPΔVif plasmid, 200 ng of VSV-g and 150 ng of either APOBEC3G or APOBEC3F plasmid. 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.

Twenty-four hours prior to infection, 3.0 x 10^5 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 2000 RPM 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 2000 RPM 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).

3.3.2 Western Blotting

Forty-eight hours post-transfection, HEK 293T producer cells were harvested and washed in cold 1X phosphate buffered saline (PBS) prior to being lysed with RIPA lysis buffer (50mM Tris pH 8.0, 150mM NaCl, 1% NP-40, 0.2% Sodium Dodecyl Sulphate (SDS), 0.5% Na-Deoxycholate and 1mM EDTA) supplemented with 1X cOmplete, EDTA-free, protease inhibitor cocktail (Roche Diagnostics Corp., Indianapolis, Indiana, USA). The cell lysates were syringed and incubated on ice for 30 minutes before being spun at 17000 x g at 4 °C in a Sorvall Legend Micro 17R microcentrifuge (Thermo Scientific) to remove any insoluble material. The cleared lysates were transferred to new 1.5 mL microcentrifuge tubes (Eppendorf Canada, Mississauga, Ontario, Canada) and a portion of each lysate was mixed with 5X

Leammli loading buffer (0.5M Tris pH 6.8, 0.2% (v/v) glycerol, 20% (w/v) SDS, 0.16% (w/v) bromophenol blue) supplemented with 5% β-mercaptoethanol (Sigma-Aldrich) at a 3:1 lysate:loading buffer ratio and boiled for 5 minutes at 100 °C. The lysates were incubated on ice for 2 minutes prior to loading on 10% SDS-polyacrylamide gel electrophoresis (PAGE) acrylamide gels (Bio-Rad Laboratories (Canada) Ltd., Mississauga, Ontario, Canada) and run at 175 V for 1 hour in 1X running buffer (190mM glycine, 0.1% (w/v) SDS, 25mM Tris-base). After the proteins were resolved, the samples were transferred at 0.35 mA for 1 hour in 1X Novex transfer buffer (190mM glycine, 25mM Tris-base, 20% (v/v) methanol) onto 0.45 μm pore Immobilon polyvinylidene fluoride (PVDF) membranes (EMD Millipore). The membranes were briefly washed twice in 1X PBS and blocked with 5% skim milk in 1X PBS-T (0.1% Tween 20) (Acros Organics, Fair Lawn, New Jersey, USA) for 1 hour at room temperature on an orbital shaker. The membranes were quickly rinsed in 1X PBS-T prior to probing with antibodies. Monoclonal anti-FLAG M2-peroxidase horse radish peroxidase (HRP)-conjugated antibody (Sigma-Aldrich) at a 1:3000 dilution was used to blot the FLAG-tagged APOBEC proteins. Primary anti-eGFP antibody (Clontech) at a 1:3000 dilution was used to blot the eGFP-tagged APOBEC proteins, with an anti-mouse secondary antibody (Cell Signaling Technology, Inc., Danvers, Massachusetts, USA) at a 1:3000 dilution. Primary anti-p24 antibody (Abcam Inc., Cambridge, England, UK) at a 1:2000 dilution was used to blot HIV-1 p24 capsid protein, with the anti-mouse secondary antibody. Subsequently, polyclonal anti-β-tubulin HRP-conjugated antibody (Abcam) at a 1:20,000 dilution was used as a loading control. All antibody dilutions were prepared in 5% skim milk in 1X PBS-T from a 1 mg/mL antibody stock. The membranes were blotted for 1 hour at room temperature or overnight at 4°C on a shaker (primary antibody only) and then washed three times for 10 minutes in 1X PBS-T and quickly rinsed in 1X PBS prior to detection. The proteins were detected using the

Clarity Western ECL Substrate kit (Bio-Rad) and exposed on CL-XPosure imaging films (Thermo Scientific) or using the ImageQuant LAS 4000 digital imaging system (GE Healthcare, Buckinghamshire, England).

3.3.3 Polymerase Chain Reaction (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 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 reaction volume of 50 μL with (final) 1U PrimeSTAR HS DNA polymerase (TaKaRa), 1X PrimeSTAR buffer (Mg^{2+}), 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 (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 (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 Sigma-Aldrich. 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.

3.3.4 Cloning

Figure A1 in Appendix

The 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 Clean-Up 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, 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.

3.3.5 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, Mg^{2+}, 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 (Table A1), and topped up to 20 µL with sterile water. The reaction plate was briefly spun down 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 quantitative 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).

3.3.6 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-deltaE-EGFPΔVif LTR (5' or 3' as applicable) sequence as a reference.

3.4 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 10^5 cells per well of a 24-well plate in 0.5 mL of 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.

3.5 Mutation Analysis of HIV-1 Patient Sequences in Database

HIV-1 patient sequences for both the 5' and 3' LTR regions were downloaded from the HIV Database (http://www.hiv.lanl.gov/content/index) on September 21st, 2015 using their Sequence Search Interface. 5' LTR sequences were defined as nucleotides 1-633 and the 3' LTR sequences were defined as nucleotides 9086-9719. Downloaded sequences were subject to strip gap handling. Each subtype was aligned separately (except for indicated CRF subtypes which were grouped) using the contig automatic assembly feature from Sequencher 4.10.1 software. All were processed using the default assembly parameters (Minimum Match Percentage-85%; Minimum Overlap-20) except for the group of CRF-08-12 (which used a Minimum Match Percentage of 70%).

3.6 Generation of Reconstituted Heavily Mutated Clones

Figure A2 in Appendix

3.6.1 HIV-1 Vector Backbone

A fragment of the viral vector pNL4-3-deltaE-EGFPΔVif was PCR amplified using (final) 2.5U Bestaq DNA polymerase (Applied Biological Materials Inc. (abm), Richmond, British Columbia, Canada), 1X PCR buffer (Mg^{2+}), 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-deltaE-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 in Section 3.3.4. 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 (Table A1).

3.6.2 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 (peGFP-C3 SacII 3' LTR FWD deg and peGFP-C3 NaeI 3' LTR REV deg (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 (Table A1) using each heavily mutated LTR from the eGFP reporter constructs as references for the reconstituted clones.

3.7 Generation of Reconstituted Revertant Clones

Figure A3 in Appendix

Site-directed mutagenesis was performed on Clone D9 of the LTR-eGFP reporter library to individually revert nine of its G-to-A mutations back to wildtype guanine residues. SDM was performed on 10-, 50-, and 100 ng of plasmid DNA template with (final) 5U Bestaq DNA polymerase, 1X PCR buffer (Mg^{2+}), 200 μM dNTP mix, 125 ng of each of the forward and

reverse primers (Table A1), 3% dimethyl sulfoxide (DMSO) (Sigma-Aldrich), and topped up to 50 uL with sterile water. The PCR conditions were as follows: an initial denaturation at 95 °C for 1 minute, followed by 18 cycles of denaturation at 95 °C for 50 seconds, annealing at 58 °C for 50 seconds, extension at 68 °C for 5 minutes, and a final extension at 68 °C for 7 minutes. Following the SDM reaction, the PCR products were digested with the enzyme *DpnI* (NEB) for 1.5 hours to chew up any methylated template DNA, leaving only the mutated unmethylated PCR product behind. The SDM reactions were then further cleaned up through isopropanol precipitation to concentrate the DNA. Each SDM reaction was transferred from its current PCR tube to a 1.5 mL microcentrifuge tube and neutralized with 5 µL (1:10) of 3M sodium acetate pH 5.2 before being topped up with 1 mL of cold 100% isopropanol. The tubes were mixed and incubated on dry ice for 30 minutes or overnight at -80 °C. The samples were then centrifuged at 17000 x g at 4 °C for 30 minutes to pellet the DNA. The isopropanol was decanted and the samples were washed with 1 mL of cold 70% ethanol to remove any salts. The ethanol was decanted and any residual ethanol was pipetted off before being air dried for 10 minutes. The DNA pellet was then resuspended in 10 µL of sterile water and rehydrated at 65 °C for 10 minutes. After cooling to room temperature, the SDMs were transformed as previously described and plated onto LB-Agar+AMP plates. Minipreps were grown and processed as similarly described in Section 3.6 and sent for Sanger sequencing using the sequencing primers from Section 3.3.6. The rest of the protocol to reconstitute the reverted LTRs follows the same instructions as in Section 3.6.2 but using Clone D9 from the LTR-eGFP reporter library as the sequencing reference.

3.8 Generation of Reconstituted Single/Double Point Mutants

Figure A4 in Appendix

The viral vector pNL4-3-deltaE-EGFPΔVif was digested with the restriction enzymes *XhoI* (sticky) and *NaeI* (blunt) to separate the 3' LTR and the rest of the viral vector. Both fragments were retained for downstream cloning. The vector pBlueScript II SK (+) (pBSK) (Stratagene, La Jolla, California, USA) was digested with the restriction enzymes *XhoI* (sticky) and *EcoRV* (blunt) (NEB) and dephosphorylated. The 3' LTR fragment from the viral vector was phosphorylated and ligated into the pBSK vector for subcloning. The ligations were transformed as previously described and plated onto LB-Agar+AMP plates previously spread with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (Sigma-Aldrich) and isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich) for blue-white screening. The positive recombinants (white colonies) were processed and sent for Sanger sequencing using the primers Seq HIV 3' LTR SDM FWD and Seq HIV 3' LTR SDM REV (Table A1). SDM was performed as similarly described in Section 3.7 on this intermediate vector to introduce single and double G-to-A point mutations into the viral 3' LTR (Table A1). In the case of the double mutants, the point mutations were introduced sequentially (ie. after confirmation of successful mutation of the first SDM, the second SDM would be performed on that same clone). The clones were transformed, processed, and sent for Sanger sequencing using the same primers as above. The successfully mutated SDM subclones were digested with the restriction enzymes *XhoI* (sticky) and *SmaI* (blunt) (NEB) and phosphorylated. The large fragment from the vector pNL4-3-deltaE-EGFPΔVif that was digested earlier was dephosphorylated and used for the reconstitution of the mutated LTRs from the subclones. The ligations were transformed onto LB-Agar+AMP plates and processed as described previously.

Reconstituted SDM clones were sent for Sanger sequencing with the same sequencing primers used throughout this section.

3.9 Stimulation Assays with Reconstituted Clones

HEK 293T producer cells were seeded at 1.0 x 10^5 cells per well of a 24-well plate in 0.5 mL of DMEM 24 hours prior to transfection until they reached approximately 80% confluency. The following day, the media of the producer cells was replaced with 1 mL of fresh DMEM and the cells were transfected with 100 ng of each reconstituted clone (heavily mutated, revertant, SDM) and 60 ng of VSV-g using GeneJuice transfection reagent according to manufacturer's specifications. The cells were incubated for 48 hours to allow for virus production. 24 hours prior to infection, HEK 293T target cells were seeded at 1.0 x 10^5 cells per well of a 24-well plate in 0.5 mL of DMEM until they reached approximately 50-60% confluency. 48 hours post-transfection, the viral supernatants from the producer cells were harvested and spun down at 2000 RPM for 10 minutes in a Sorvall centrifuge to pellet any cell debris. Prior to infection, the media of the target cells was replaced with 1 mL of fresh DMEM plus polybrene. The target cells were then infected with equal volumes of viral supernatant (two wells per clone) and spinoculated at 2000 RPM for 1 hour. 24 hours post-infection, the media of one set of infections was replaced with 1 mL of fresh complete DMEM plus 10 ng/µL of phorbol myristate acetate (PMA) (Sigma-Aldrich) and 1 µM ionomycin (Sigma-Aldrich) for stimulation. The media of the other set of infections was replaced with 1 mL of fresh complete DMEM alone. 24 hours after media replacement/stimulation, the cells were harvested in fluorescence-activated cell sorting (FACS) sort buffer (1X PBS Ca/Mg^{2+} free, 1mM EDTA, 25mM HEPES pH 7.0, 1% FBS heat inactivated), fixed with 4% paraformaldehyde (PFA), and measured for eGFP expression and fluorescence intensity by

flow cytometry on a BD LSRFortessa or BD FACSCelesta flow cytometer (BD Biosciences, Mississauga, Ontario, Canada).

3.10 Generation of U3 mut/R-U5 wt Clones

Figure A5 in Appendix

The APOBEC3-mutated U3 region of each of the heavily mutated constructs was cloned into a peGFP-C3 expression plasmid containing wildtype R and U5 regions.

3.10.1 Wildtype R and U5 Regions from pNL4-3-deltaE-EGFPΔVif

A semi-nested PCR was performed to amplify the R and U5 regions of the 3' LTR of pNL4-3-deltaE-EGFPΔVif. The PCR reagents and conditions were the same as those described in Section 3.3.3 except for the extension time of 30 seconds and the second-round primers used: HIV 3' LTR R-U5 wt FWD and HIV 3' LTR R-U5 wt REV (Table A1). The second-round forward primer included an additional restriction endonuclease site (*NotI*) for downstream cloning. The PCR product was digested with the restriction enzymes *AseI* and *NheI* and phosphorylated.

3.10.2 Intermediate peGFP-C3 Plasmid Containing Wildtype R and U5 Regions

The CMV promoter of the peGFP-C3 expression plasmid was cut out using the restriction enzymes *AseI* and *NheI* and the leftover vector was dephosphorylated. The amplicon from the previous step above was ligated into this expression plasmid and transformed as previously described onto LB-Agar+KAN plates. Minipreps were grown and processed as described previously and sent for Sanger sequencing to confirm successful insertion. The plasmid was then further digested with the restriction enzymes *AseI* and *NotI* (NEB) and dephosphorylated.

3.10.3 Mutated U3 Regions from LTR-eGFP Constructs

The U3 region of each heavily mutated LTR-eGFP reporter construct was PCR amplified using the primers peGFP-C3 U3 mut FWD and peGFP-C3 U3 mut REV (Table A1) and the same PCR reagents and conditions as above except for an extension time of 45 seconds. Each of the amplicons was digested with the restriction enzymes *AseI* and *NotI* and phosphorylated. The amplicons were then ligated into the intermediate plasmid created above and transformed as previously described onto LB-Agar+KAN plates. The clonings were processed as described previously and sent for Sanger sequencing using the same sequencing primers as in Section 3.3.6 using each heavily mutated LTR as reference sequences.

3.11 Functional Assays with U3 mut/R-U5 wt Clones

The protocol follows the same procedures outlined in Section 3.4 except that the HcRed-expressing Tat plasmid was used. The samples were analyzed on a BD LSRFortessa or BD FACSCelesta flow cytometer. The data obtained from these clones was compared to those obtained in the original LTR-eGFP reporter library.

3.12 Statistical Analysis

All functional, stimulation, infectivity and restriction assays were performed at least three times from at least three independent transfections and the results were presented as the mean relative to the wildtype control plus standard error of the mean (SEM) or standard deviation (SD). All statistical analyses were performed using GraphPad Prism 7 software (GraphPad Software, La Jolla, California, USA).

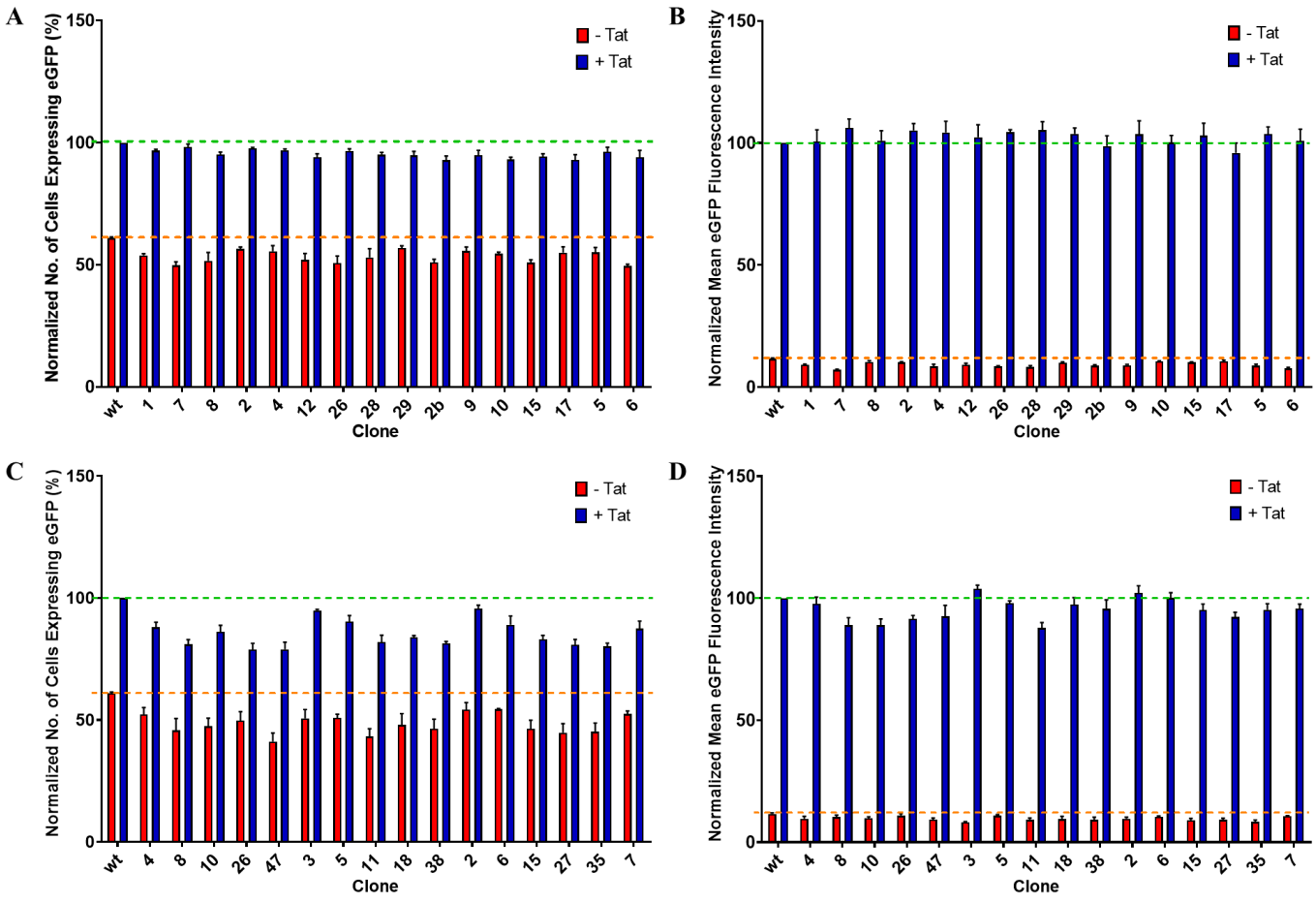
4.0 RESULTS

4.1 Mutations in the 3' LTR (not 5' LTR) Influence HIV-1 Promoter Functionality

In order to investigate the effects of APOBEC3 deamination on the activity of the HIV-1 promoter, the 5' LTR of HIV-1 proviruses produced in the presence of A3G or A3F was cloned upstream of an eGFP reporter gene. Out of 83 clones sent for Sanger sequencing for G-to-A hypermutation, 73 returned mutated, with 32 (16 clones each from 5' LTRs of HIV-1 mutated by A3G and A3F) uniquely mutated and deposited into our library. Each clone returned with between 1 to 4 G-to-A mutations at five different positions in the 5' LTR (positions 2, 6, 7, 16, and 620, if we designate the first base of the LTR as position 1). In order to test the effects of these mutations on the activity of the promoter, a functional transfection assay was performed with the library of 5' LTR-eGFP reporter constructs (Figure 4.1). Each clone was transfected in the absence and presence of the viral regulatory protein Tat (as Tat transactivation is important for efficient HIV-1 transcription) to visualize both basal as well as Tat-inducible transcription, respectively. Promoter functionality was assessed by measuring the expression of the eGFP reporter gene under the control of the A3-mutated promoters. The results were normalized to a wildtype control that was not exposed to A3 (100% eGFP expression was normalized to the wildtype control + Tat). Half of the cells transfected with the wildtype control expressed the eGFP reporter gene in the absence of Tat, demonstrating that a basal level of transcription was detectable from the promoters. With the addition of Tat, the percentage of cells that expressed the eGFP reporter gene increased, confirming the functionality of the Tat plasmid in our system as well as the protein's importance in viral transcription activation. At the basal level of transcription (without Tat), the mean level of eGFP fluorescence was quite low, around 10% for the wildtype control. With Tat induction, the mean fluorescence intensity (MFI) of eGFP dramatically increased. All clones in both the

A3G and A3F sets behaved similarly to the wildtype control. From this, we gathered that the few mutations generated in the 5' LTR by the A3 proteins had a negligible effect on promoter activity. When looking at the positions of the mutations that were induced, they were located at the immediate beginning and end of the 5' LTR. Upon further inspection, it became apparent that the mutations that were obtained were actually an artifact of the G/A; C/T degenerate primers designed to amplify the 5' LTR. Therefore, surprisingly, the 5' LTR was not mutated by the APOBEC3 proteins in this single-round infection system. At this point, it would appear that the A3s were not targeting the HIV-1 promoter for deamination. We also conclude that the putative deamination target sites that were mutated because of the degenerate primers had no impact on promoter activity.

Figure 4.1. Functional transfection assays with 5' LTR-eGFP reporter constructs. The 5' LTR of HIV-1 proviruses produced in the presence of A3G (*A* and *B*) or A3F (*C* and *D*) was cloned upstream of an eGFP reporter gene. HEK 293T cells were transfected with each clone in the presence and absence of the viral protein, Tat. 48 hours post-transfection, the functionality of the promoter was analyzed using flow cytometry to measure the percentage of cells expressing eGFP (*left*) as well as the mean eGFP fluorescence intensity (*right*). The results were normalized to the wildtype (wt) control that has not been exposed to A3. The orange dashed line indicates the wt standard in the absence of Tat, while the green dashed line indicates the wt standard in the presence of Tat. Clones are ordered from least number of mutations (1) to most number of mutations (4). Three independent transfections were performed and the results are presented as mean + standard error of the mean (SEM).

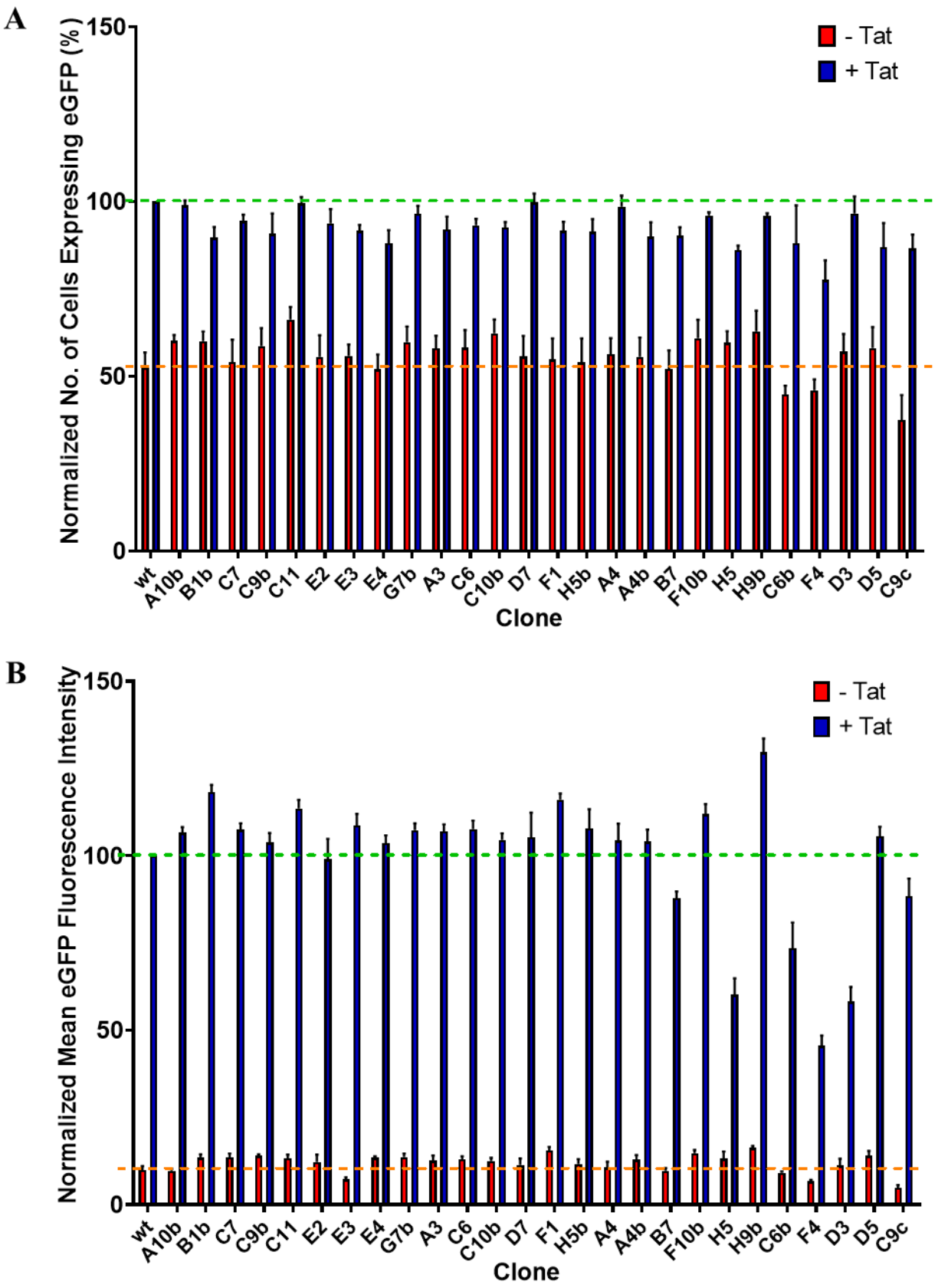

A
Normalized No. of Cells Expressing eGFP (%)
150
100
50
0
- Tat
+ Tat
wt 1 7 8 2 4 12 26 28 29 2b 9 10 15 17 5 6
Clone
B
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
- Tat
+ Tat
wt 1 7 8 2 4 12 26 28 29 2b 9 10 15 17 5 6
Clone
C
Normalized No. of Cells Expressing eGFP (%)
150
100
50
0
- Tat
+ Tat
wt 4 8 10 26 47 3 5 11 18 38 2 6 15 27 35 7
Clone
D
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
- Tat
+ Tat
wt 4 8 10 26 47 3 5 11 18 38 2 6 15 27 35 7
Clone

However, it is known in the literature that there is a gradient of mutations from the 5' end (less mutated) to the 3' end (more mutated) of the viral genome due to the processivity of the APOBEC3 proteins (126, 139, 180). During reverse transcription, the 3' end of the viral DNA remains single-stranded longer than the 5' end, leading to the 5'→3' intensifying gradient of APOBEC3-induced mutations observed in the viral genome. Therefore, we postulated that perhaps there may be more mutations generated in the 3' LTR of the virus. During reverse transcription when strand transfer occurs, the mutations in the 3' LTR would be copied over to the 5' end of the genome and influence promoter activity. Thus, the above experiments were repeated with the 3' LTR. After cloning the 3' LTR of HIV-1 proviruses produced in the presence of A3G and A3F into the eGFP expression vector, the identification of mutated clones was facilitated by utilizing a technique called high resolution melt (HRM). HRM is a technique used to quickly detect transition mutations in double-stranded DNA (174). It incorporates an intercalating dye, such as SYBR Green, which binds only to double-stranded DNA and fluoresces. The process starts with PCR to amplify the region of interest (in our case, the LTRs of the clones). Afterwards, there is a melting process whereby the amplicon is gradually heated and the level of fluorescence is measured in real-time. Eventually, a temperature will be reached where the two strands in a sample melt apart and the level of SYBR Green fluorescence drops. Because the A3s cause G-to-A hypermutation, mutated samples would be more AT-rich than less mutated ones and would require lower temperatures to melt than the control. The control in this case is HIV-1 produced in the presence of APOBEC2 (A2), a member of the APOBEC family known to have no deamination activity (124). The results of the HRM would then be plotted as aligned melt curves. The samples with melt curves to the left of the control (ie. melted at lower temperatures than the non-mutating

A2 control) would be the ones we select to send for sequencing as they are the most likely to contain mutations.

Out of 94 3' LTR HIV-1 A3F clones sent for Sanger sequencing, 56 returned mutated, with 26 uniquely mutated and deposited into our library. Each clone returned with between 1 to 7 G-to-A mutations at 32 different positions along the 3' LTR. Unlike the 5' LTR clones, the mutations in the 3' LTR were found along its entire length. A functional transfection assay was then performed with the library of 3' LTR HIV-1 A3F clones (Figure 4.2). Similar to the wildtype control, in the absence of Tat, half of the cells for each of the clones expressed the eGFP reporter gene. With the addition of Tat, the percentage of cells expressing eGFP in the clones increased similarly to the wildtype control. It is only when we look at the mean eGFP fluorescence intensity of the clones that we are able to observe the slight modulation that some of the A3F-induced mutations have on the activity of the promoter. At the basal level of transcription, the mean level of eGFP fluorescence was once again quite low, around 10% for all of the clones, with some slightly higher and lower than the wildtype control (clones F1, F10b, H9b and E3, F4, C9c, respectively). With Tat induction, the MFI of eGFP greatly increased, but not to the same extent for all clones. Clones H5, B7, F4, C6b, D3, and C9c, for example, did not induce to quite the same extent as some of the other clones or the wildtype control. Furthermore, there were some clones that possessed mutations that appeared to allow them to reach an even higher level of fluorescence than the wildtype (clones C11, B1b, F1, F10b, H9b).

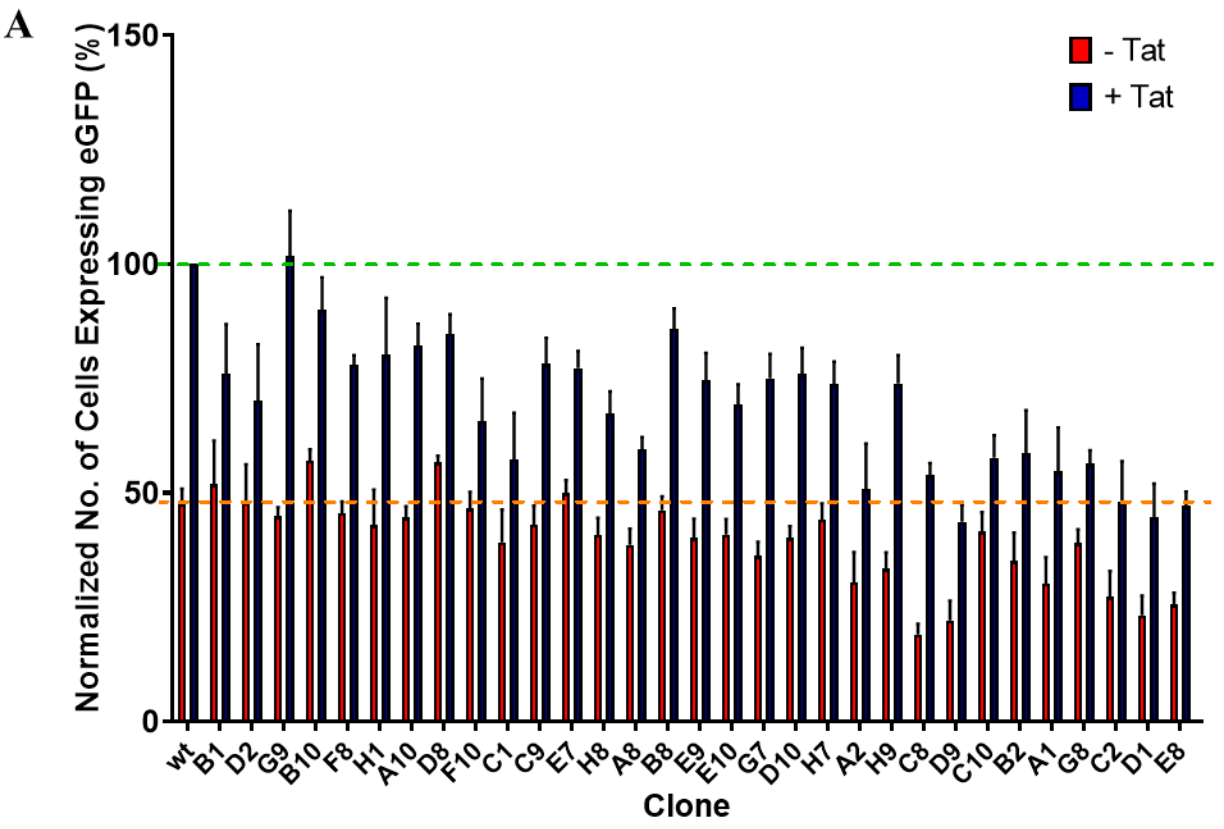
Figure 4.2. Functional transfection assay with 3' LTR-eGFP HIV-1 A3F reporter constructs. The 3' LTR of HIV-1 proviruses produced in the presence of A3F was cloned upstream of an eGFP reporter gene. HEK 293T cells were transfected with each clone in the presence and absence of the viral protein, Tat. 48 hours post-transfection, the functionality of the promoter was analyzed using flow cytometry to measure the percentage of cells expressing eGFP (*A*) as well as the mean eGFP fluorescence intensity (*B*). The results were normalized to the wildtype (wt) control that has not been exposed to A3. The orange dashed line indicates the wt standard in the absence of Tat, while the green dashed line indicates the wt standard in the presence of Tat. Clones are ordered from least number of mutations (1) to most number of mutations (7). Five independent transfections were performed and the results are presented as mean + standard error of the mean (SEM).

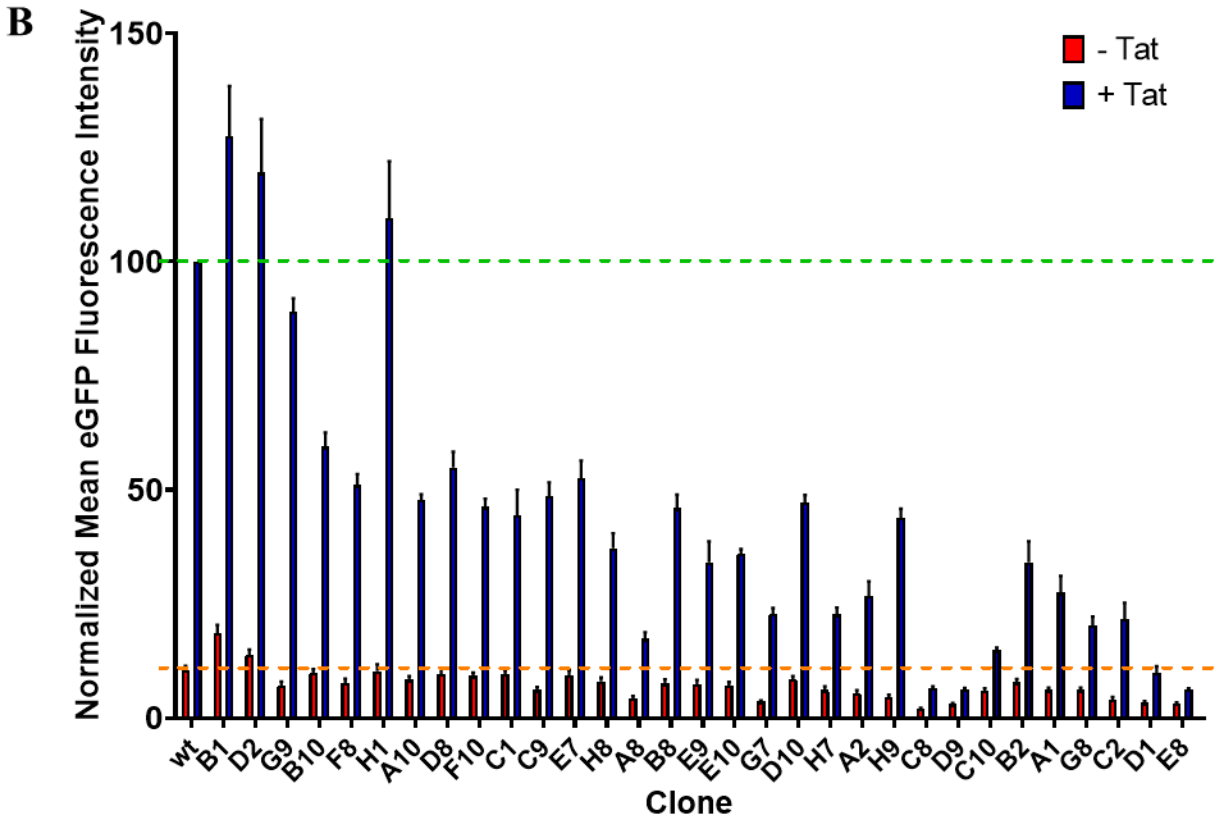

A
Normalized No. of Cells Expressing eGFP (%)
150
100
50
0
- Tat
+ Tat
wt
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
C6
C10b
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone
B
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
- Tat
+ Tat
wt
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
C6
C10b
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone

Out of 31 3' LTR HIV-1 A3G clones sent for Sanger sequencing, all 31 returned uniquely mutated and were deposited into our library. Each clone returned with between 2 to 37 G-to-A mutations at 65 different positions along the 3' LTR. Once again, unlike the 5' LTR clones and similar to the A3F set of 3' LTR clones, the A3G-induced mutations were found along the entire length of the 3' LTR. It was striking to see how abundantly mutated the 3' LTR was by A3G. To investigate whether these numerous mutations had an effect on promoter activity, a functional transfection assay was performed using the library of 3' LTR HIV-1 A3G clones (Figure 4.3). We discovered that the mutations caused by A3G on the 3' LTR had a wide effect on the activity of the promoter. Unlike in the previous experiments where without Tat half of the cells in the clones were still able to express the eGFP reporter gene, with the A3G-induced mutations, the basal transcription capability of some of the clones was compromised when compared to the wildtype control. Most notably, the cells transfected with clones A2, H9, D9, C8, A1, C2, D1, and E8 were not able to express the eGFP reporter gene as well as the wildtype control in the absence of Tat (about a third to one half less cells expressing eGFP). With the addition of Tat, the percentage of cells able to express eGFP in the clones increased, but not to the same extent as the wildtype control. In fact, most of the clones expressed eGFP only marginally better with Tat induction. Clones A2, D9, C2, D1, and E8 were not able to reach even wildtype basal levels of eGFP expression. However, what is more significant is the level at which these cells fluoresce. In the absence of Tat, as with the previous experiments, there was a basal level of transcription with a mean level of eGFP fluorescence that was quite low. The mutations caused by A3G were able to affect the basal transcription of the clones and most clones had a reduced mean level of eGFP fluorescence compared to the wildtype control. The basal transcription of a few clones appeared to increase slightly with the A3G-induced mutations that they possessed (ie. clones B1, D2), while others

were dramatically reduced (ie. clones A8, G7, A2, H9, D9, C8, C2, D1, E8). With Tat induction, the mean level of eGFP fluorescence increased to varying degrees. A few clones were induced to higher levels than the wildtype control (ie. clones B1, D2, H1). However, the majority of clones did not reach induction levels as high as the wildtype and were impaired by 50% or more. Some clones were no longer responsive to Tat-induction with the mutations that they acquired and could not reach even wildtype levels of basal transcription (ie. clones D9, C8, D1, E8). These promoters were inactivated by their mutation burden. Still, there were clones that carried mutations that substantially weakened their transcription capability, yet were still Tat-inducible (ex. clones A8, G7, H7, A2, G8, A1, C2). These clones are of potential interest in searching for latency-inducing mutations. One should also note that some clones that had fewer cells expressing eGFP did not necessarily fluoresce the dimmest and vice versa. For example, clones B1 and D2 fluoresced brighter than the wildtype control despite having less cells expressing eGFP. Similarly, clones G9 and B10 had as many cells expressing eGFP as the wildtype control yet did not fluoresce as brightly. This is because the percentage of cells expressing eGFP represents the percentage of clones *induced* to transcribe, while the mean fluorescence intensity represents the *level* at which these clones transcribe, which can be exclusive entities. For most of the clones though, the percentage of cells expressing eGFP did track well with the intensity of fluorescence of the cells (ie. their induction capacity correlated with their transcription capability).

Figure 4.3. Functional transfection assay with 3' LTR-eGFP HIV-1 A3G reporter constructs. The 3' LTR of HIV-1 proviruses produced in the presence of A3G was cloned upstream of an eGFP reporter gene. HEK 293T cells were transfected with each clone in the presence and absence of the viral protein, Tat. 48 hours post-transfection, the functionality of the promoter was analyzed using flow cytometry to measure the percentage of cells expressing eGFP (*A*) as well as the mean eGFP fluorescence intensity (*B*). The results were normalized to the wildtype (wt) control that has not been exposed to A3. The orange dashed line indicates the wt standard in the absence of Tat, while the green dashed line indicates the wt standard in the presence of Tat. Clones are ordered from least number of mutations (2) to most number of mutations (37). Five independent transfections were performed and the results are presented as mean + standard error of the mean (SEM).


A
Normalized No. of Cells Expressing eGFP (%)
150
100
50
0
- Tat
+ Tat
wt B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 H8 A8 B8 E9 E10 G7 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone


B
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
- Tat
+ Tat
wt B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 H8 A8 B8 E9 E10 G7 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone

4.2 Mutations are Present under Transcription Factor Binding Sites

To determine whether the mutations were affecting transcription factor binding sites (TFBS), a map of transcription factor binding sites in the HIV-1 LTR was created from sequences and positions obtained in the literature (181–185). The sites were mapped to scale above a tick plot of putative deamination target sites in the LTR. The mutations for each individual clone were then distributed in separate tick plots underneath. A combined tick plot of all sites actually targeted among the clones was then created. Separate maps were created for deamination target sites in a 5'GA context for A3F (Figure 4.4) and in a 5'GG context for A3G (Figure 4.5). 53 putative deamination target sites for A3F exist in the LTR. Among the 3' LTR HIV-1 A3F clones, 32 positions were mutated along the LTR, some of which were in a 5'GG context (72% 5'GA and 28% 5'GG). Not all 32 mutations were located at positions that affected transcription factor binding sites, with 11 different sites affected, along with the *Nef* gene and the TAR element. Although 46 putative deamination target sites exist for A3G in the LTR, 65 different positions were mutated in the 3' LTR HIV-1 A3G clones. This is because 35% of the mutations that were induced were in a 5'GA context and not all 46 putative 5'GG deamination target sites were mutated. Of these 65 positions, 16 different transcription factor binding sites were affected, along with the *Nef* gene and TAR element.

Figure 4.4. Distribution of HIV-1 LTR bases affected by A3F deamination. Schematic representation of the HIV-1 LTR with associated transcription factor binding sites, viral gene *Nef*, and TAR element. Graphical representations of the putative 5'GA deamination sites as well as the combined targeted sites are shown. Black-coloured ticks indicate mutations in a 5'GA context while red-coloured ticks indicate mutations in a 5'GG context. The total number (n) of affected bases and the proportion (%) of bases mutated in a 5'GA or 5'GG context are shown to the right of the targeted tick plot. Mutations in individual clones are displayed underneath. Clones are ordered from least number of mutations (1) to most number of mutations (7). The number of mutations in each individual clone are displayed in parentheses.

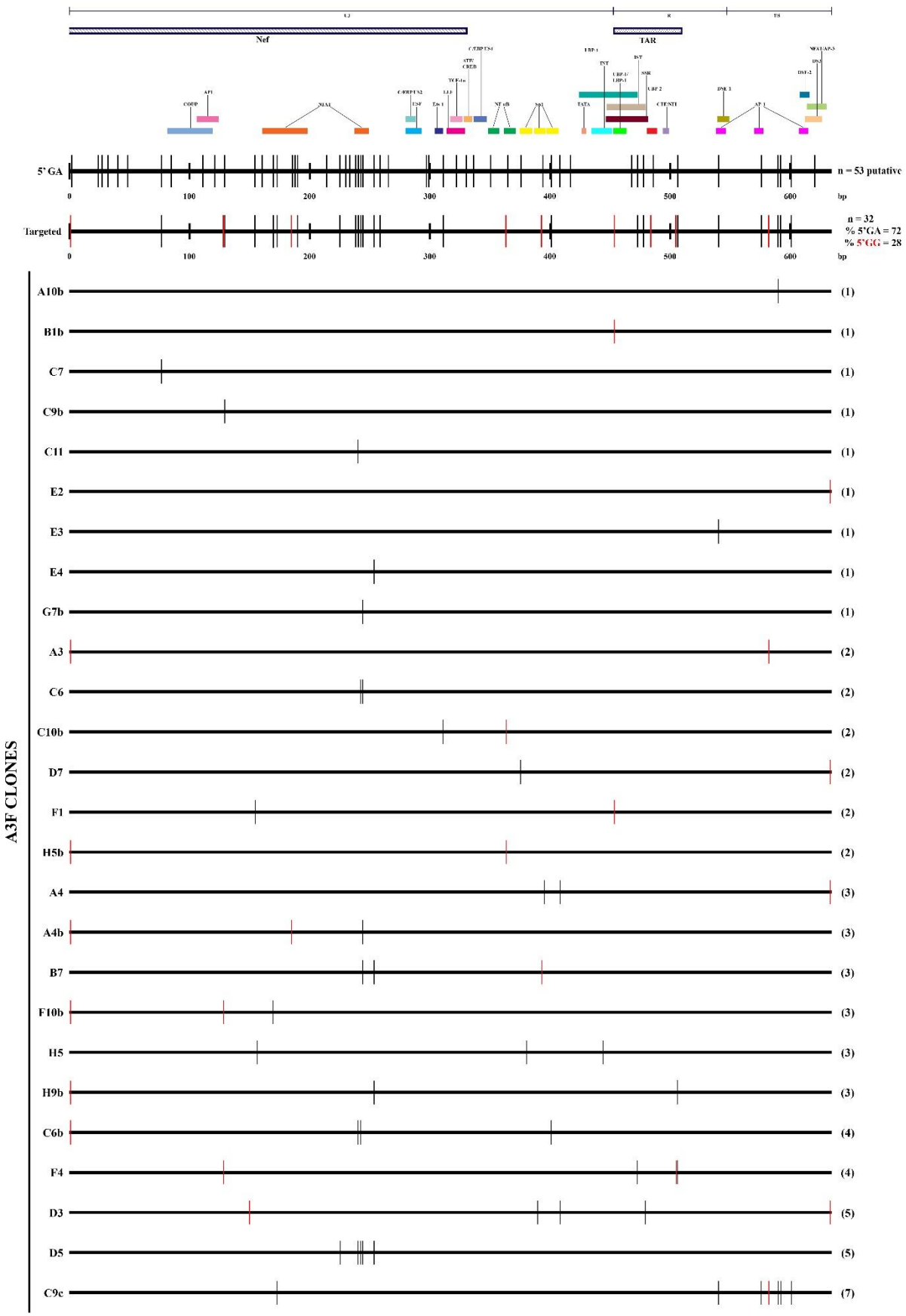
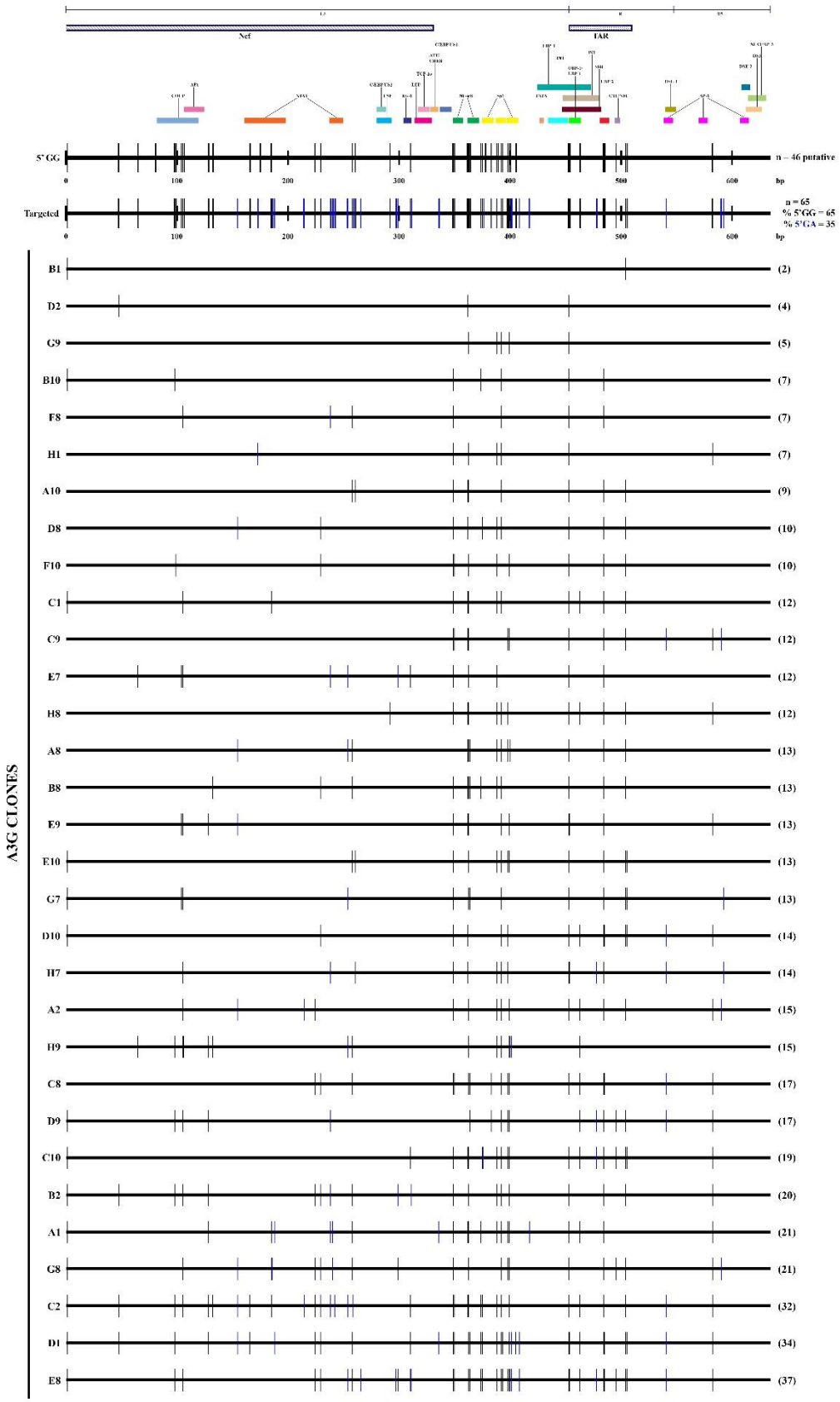

U3
R
U5
Nef
TAR
5'GA
n = 53 putative
0
100
200
300
400
500
600
bp
Targeted
n = 32
% 5'GA = 72
% 5'GG = 28
A3F CLONES
A10b (1)
B1b (1)
C7 (1)
C9b (1)
C11 (1)
E2 (1)
E3 (1)
E4 (1)
G7b (1)
A3 (2)
C6 (2)
C10b (2)
D7 (2)
F1 (2)
H5b (2)
A4 (3)
A4b (3)
B7 (3)
F10b (3)
H5 (3)
H9b (3)
C6b (4)
F4 (4)
D3 (5)
D5 (5)
C9c (7)

Figure 4.5. Distribution of HIV-1 LTR bases affected by A3G deamination. Schematic representation of the HIV-1 LTR with associated transcription factor binding sites, viral gene *Nef*, and TAR element. Graphical representations of the putative 5'GG deamination sites as well as the combined targeted sites are shown. Black-coloured ticks indicate mutations in a 5'GG context while blue-coloured ticks indicate mutations in a 5'GA context. The total number (n) of affected bases and the proportion (%) of bases mutated in a 5'GG or 5'GA context are shown to the right of the targeted tick plot. Mutations in individual clones are displayed underneath. Clones are ordered from least number of mutations (2) to most number of mutations (37). The number of mutations in each individual clone are displayed in parentheses.


Nef
TAR
5'GG
n – 46 putative
Targeted
n = 65
% 5'GG = 65
% 5'GA = 35
bp
A3G CLONES
B1 (2)
D2 (4)
G9 (5)
B10 (7)
F8 (7)
H1 (7)
A10 (9)
D8 (10)
F10 (10)
C1 (12)
C9 (12)
E7 (12)
H8 (12)
A8 (13)
B8 (13)
E9 (13)
E10 (13)
G7 (13)
D10 (14)
H7 (14)
A2 (15)
H9 (15)
C8 (17)
D9 (17)
C10 (19)
B2 (20)
A1 (21)
G8 (21)
C2 (32)
D1 (34)
E8 (37)

To determine whether there were recurring mutations that affected promoter activity, the number of G-to-A mutations per targeted position was plotted in a histogram. The bars were coloured according to the different transcription factor binding sites that were affected. Separate histogram plots were generated for the 3' LTR HIV-1 A3F clone set (Figure 4.6) and the 3' LTR HIV-1 A3G clone set (Figure 4.7). Some positions were not mutated, yet others were mutated but did not affect any transcription factor binding sites. We can also see that a mutation at one position in the LTR has the potential to affect more than one transcription factor binding site as some of the sites overlap. For the A3F set, mutations targeted positions that affected the following transcription factor binding sites: nuclear factor of activated T cells (NFAT), v-Ets erythroblastosis virus E26 oncogene homologue 1 (Ets-1), nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), specificity protein 1 (Sp1), untranslated binding protein 1/leader binding protein 1 (UBP-1/LBP-1), leader binding protein 1 (LBP-1), start site region (SSR), inducer of small transcripts (IST), untranslated binding protein 2 (UBP-2), downstream element 1 (DSE-1), activator protein 1 (AP-1), as well as the *Nef* gene and TAR element. Hot spots for deamination could not be readily identified for the A3F clones due to how sparsely each affected transcription factor binding site was hit, although NFAT may be a possibility. For the A3G set, mutations targeted positions that affected the following transcription factor binding sites: chicken ovalbumin upstream promoter (COUP), AP1, NFAT, upstream stimulatory factor (USF), Ets-1, activating transcription factor/cAMP response element binding (ATF/CREB), NF-κB, Sp1, UBP-1/LBP-1, LBP-1, SSR, IST, UBP-2, CCAAT box-binding transcription factor/nuclear factor I (CTF/NFI), DSE-1, AP-1, as well as the *Nef* gene and TAR element. There were clear peaks where certain positions were mutated more frequently. Some positions were mutated more than 20 times and were considered hot spots for A3G deamination. Mutations at the four

prominent peaks affected binding sites for NF-κB, Sp1, UBP-1/LBP-1, LBP-1, SSR, IST, UBP-2, and TAR. One should also note that NF-κB and TAR were hit frequently at two different positions in the peaks (multiple hot spots).

Figure 4.6. Histogram plot for A3F deamination hot spots in the HIV-1 LTR. The number of G-to-A mutations was plotted for each A3F-targeted position in the HIV-1 LTR. Bars are coloured according to the transcription factor binding site (TFBS) that is affected. In the case that multiple TFBSs are hit for the same position, multiple bars are displayed for that position. Bars are left empty if they do not affect any TFBS. Positions in black text indicate bases mutated in a 5'GA context while positions in red text indicate bases mutated in a 5'GG context.

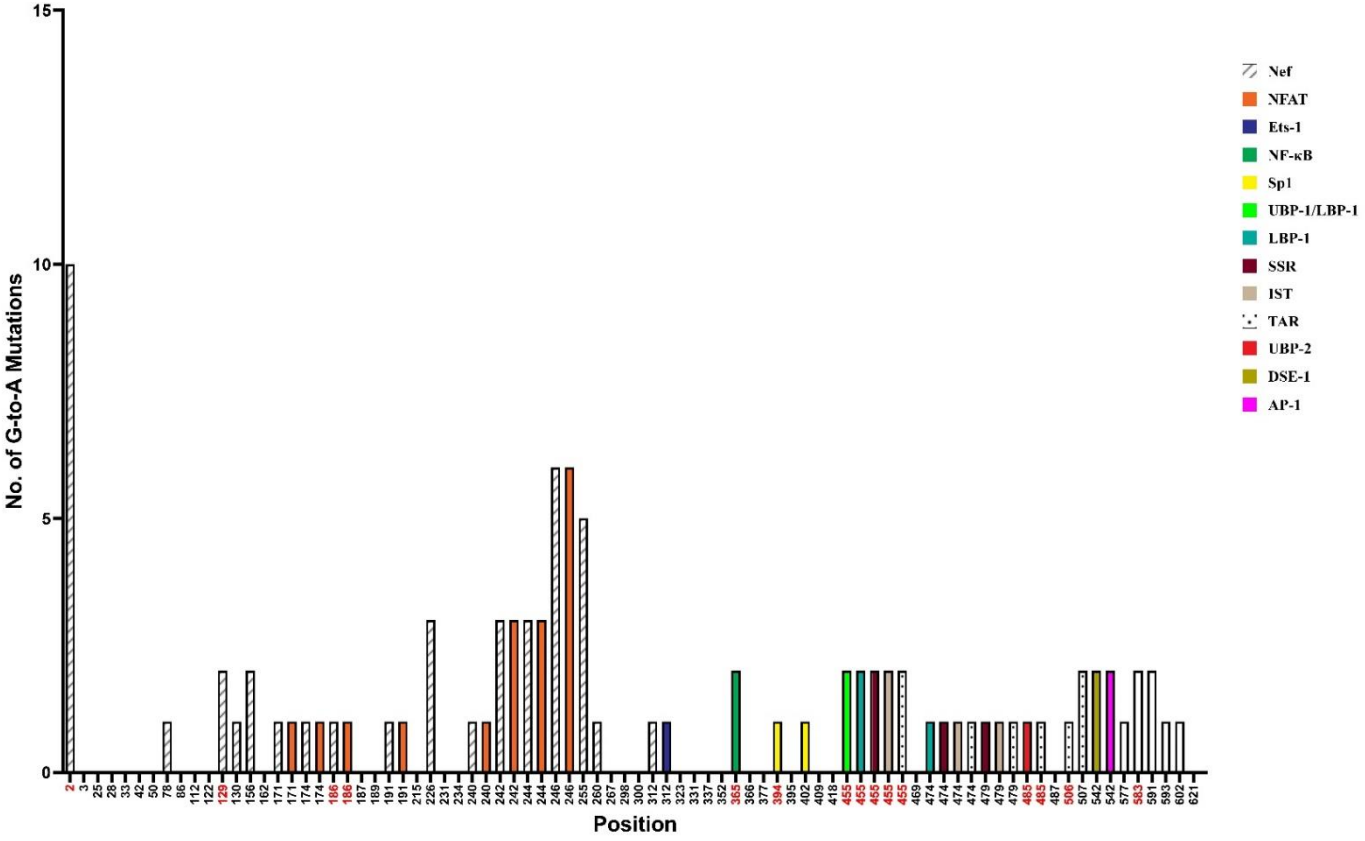
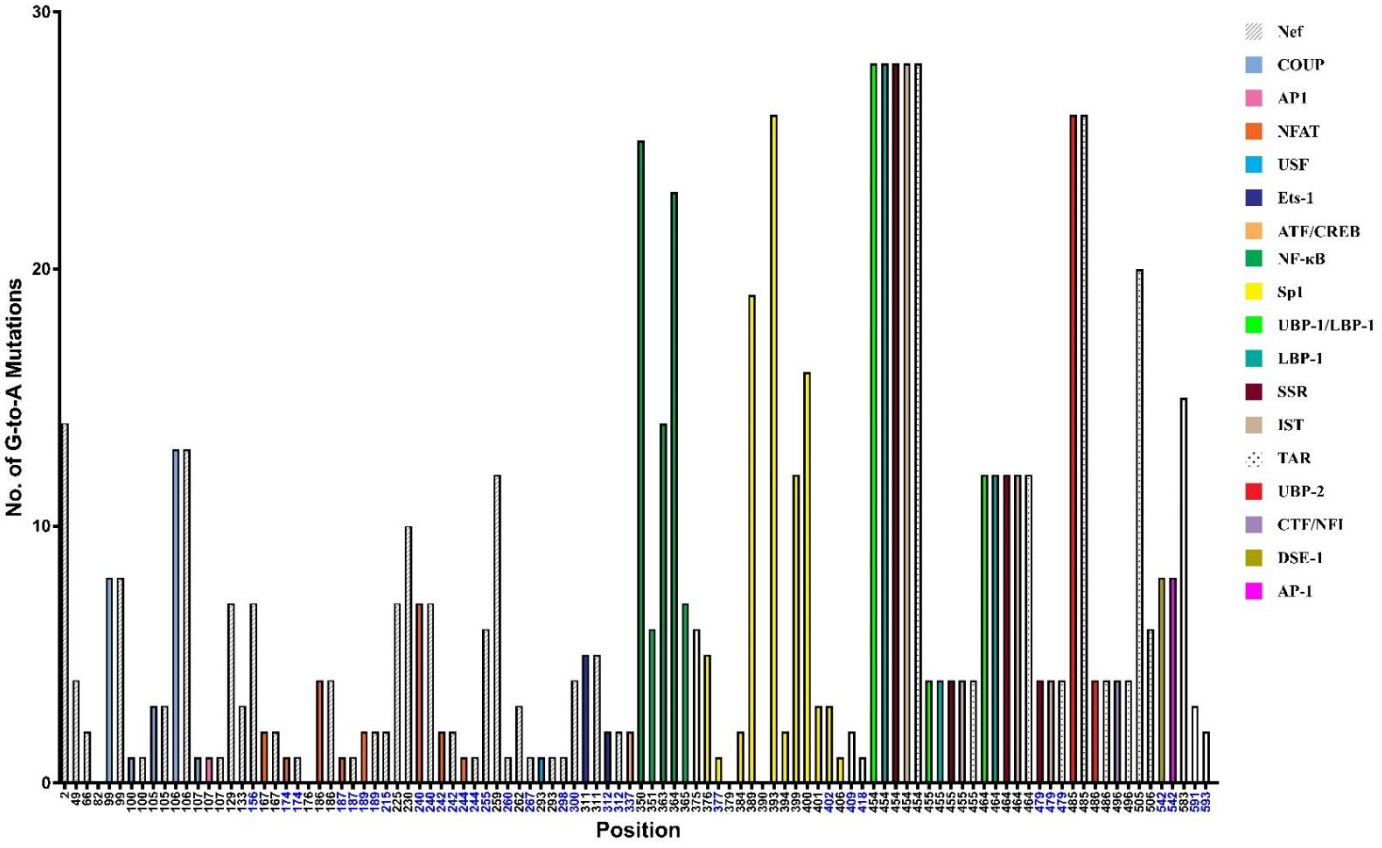

No. of G-to-A Mutations
15
10
5
0
Position
Nef
NFAT
Ets-1
NF-κB
Sp1
UBP-1/LBP-1
LBP-1
SSR
IST
TAR
UBP-2
DSE-1
AP-1

Figure 4.7. Histogram plot for A3G deamination hot spots in the HIV-1 LTR. The number of G-to-A mutations was plotted for each A3G-targeted position in the HIV-1 LTR. Bars are coloured according to the transcription factor binding site (TFBS) that is affected. In the case that multiple TFBSs are hit for the same position, multiple bars are displayed for that position. Bars are left empty if they do not affect any TFBS. Positions in black text indicate bases mutated in a 5'GG context while positions in blue text indicate bases mutated in a 5'GA context.


No. of G-to-A Mutations
30
20
10
0
Position
Nef
COUP
AP1
NFAT
USF
Ets-1
ATF/CREB
NF-κB
Sp1
UBP-1/LBP-1
LBP-1
SSR
IST
TAR
UBP-2
CTF/NFI
DSE-1
AP-1

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

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) and A3F (Figure 4.10) 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 ≤ 50 was 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 4.8. Pictogram displaying the frequency of A3-induced mutations at the indicated transcription factor binding sites in the HIV-1 LTR from heavily mutated *in vitro* clones or patient samples of indicated subtypes. HIV-1 patient sequences for both the 5' and 3' LTR regions were downloaded from the Los Alamos HIV Database (http://www.HIV.lanl.gov/content/index) using their Sequence Search Interface. The frequency of mutations at TFBSs in the HIV-1 patient sequences were compared to their mutation frequency in *in vitro* experimental clones. Circles increase in size to represent the frequency of mutation at the indicated TFBSs.

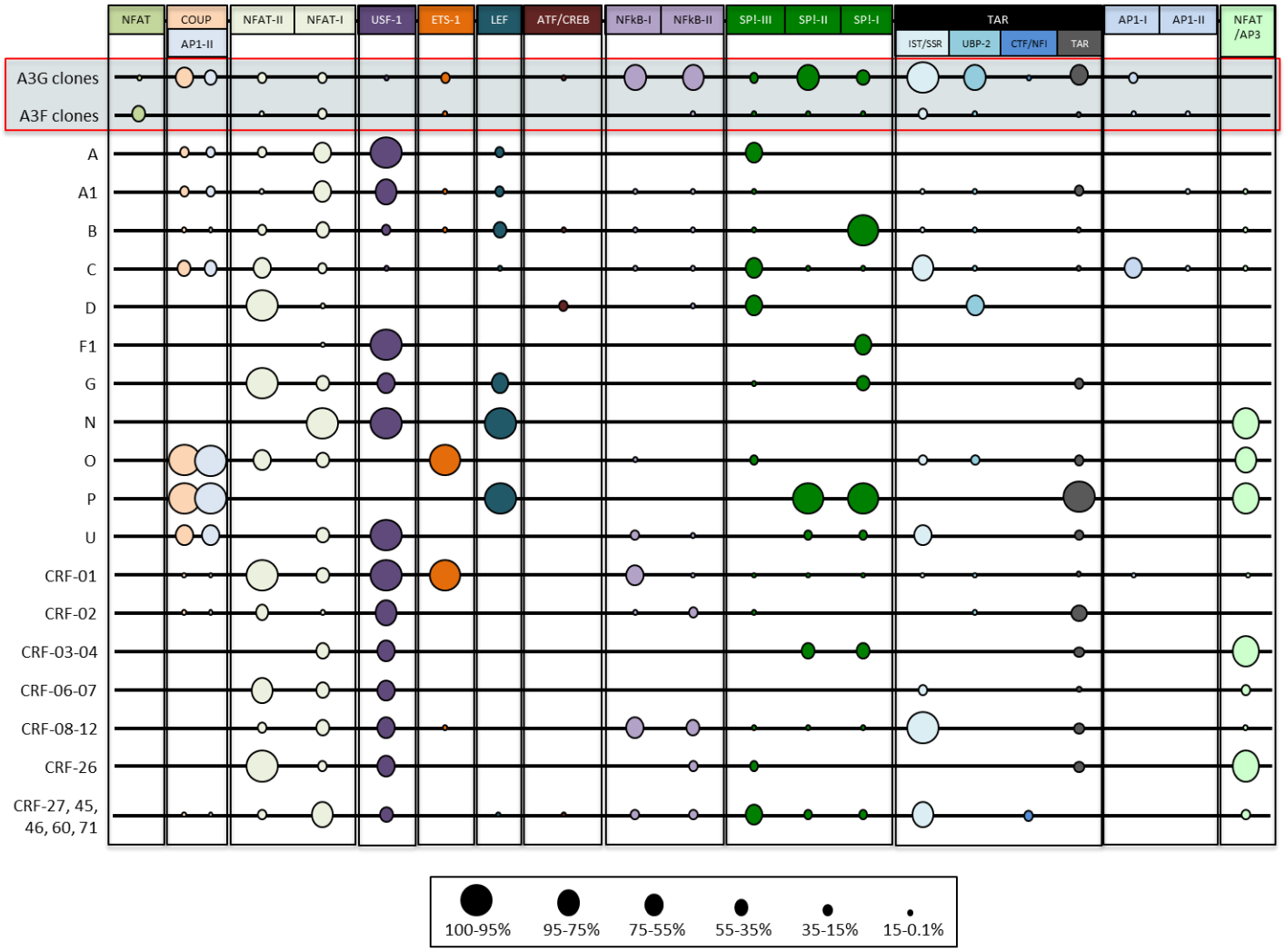

NFAT
COUP
AP1-II
NFAT-II
NFAT-I
USF-1
ETS-1
LEF
ATF/CREB
NFkB-I
NFkB-II
SP!-III
SP!-II
SP!-I
TAR
IST/SSR
UBP-2
CTF/NFI
TAR
AP1-I
AP1-II
NFAT /AP3
A3G clones
A3F clones
A
A1
B
C
D
F1
G
N
O
P
U
CRF-01
CRF-02
CRF-03-04
CRF-06-07
CRF-08-12
CRF-26
CRF-27, 45, 46, 60, 71
100-95%
95-75%
75-55%
55-35%
35-15%
15-0.1%

Figure 4.9. Box plot depicting the individual mean fluorescence intensities of the HIV-1 3' LTR A3G clones at each mutated position. The mean eGFP fluorescence intensity obtained for individual clones (dots) was plotted for each position mutated. The box plot represents the mean fluorescence intensity obtained for all clones with the specified site mutated. The location of affected transcription factor binding sites is also included.

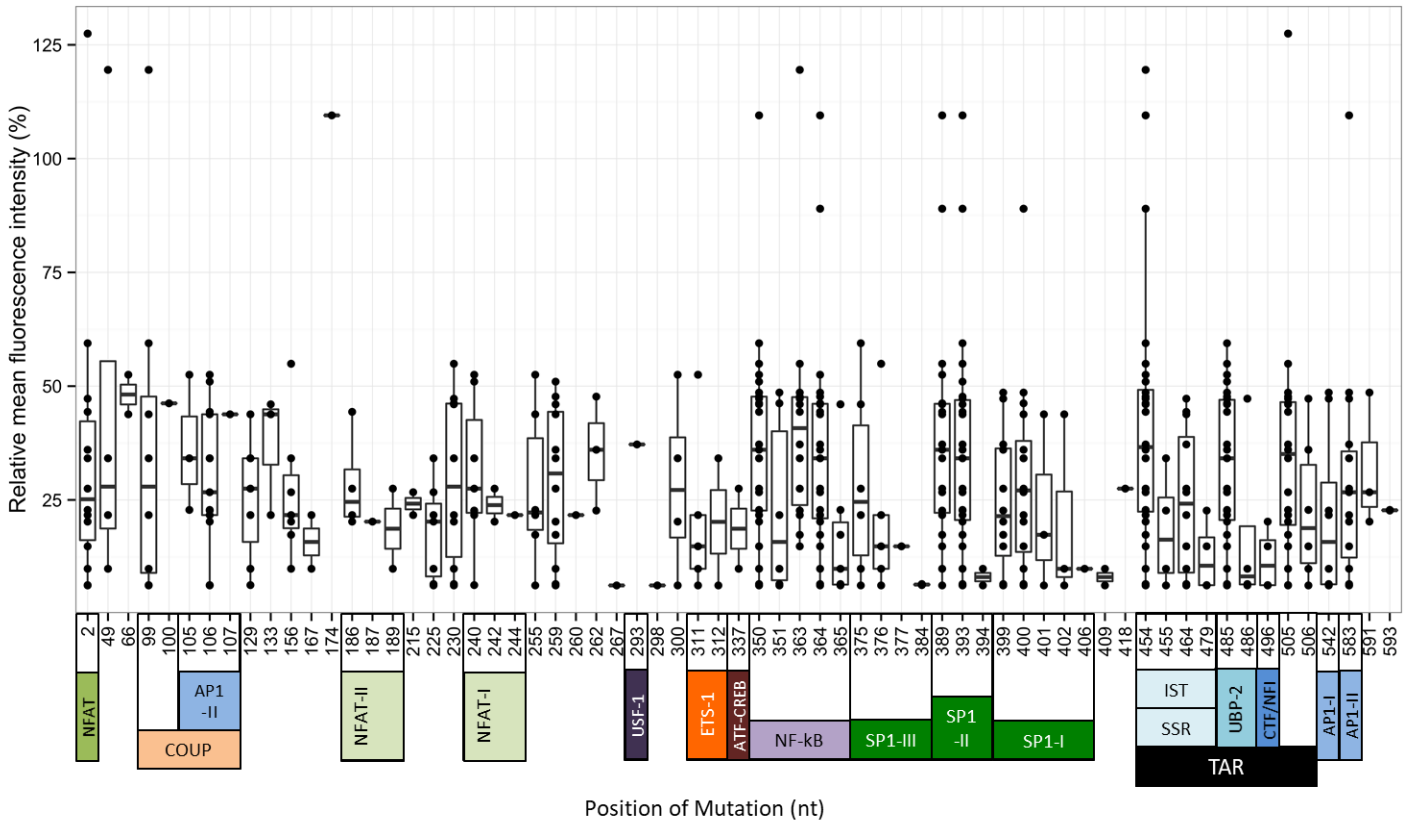
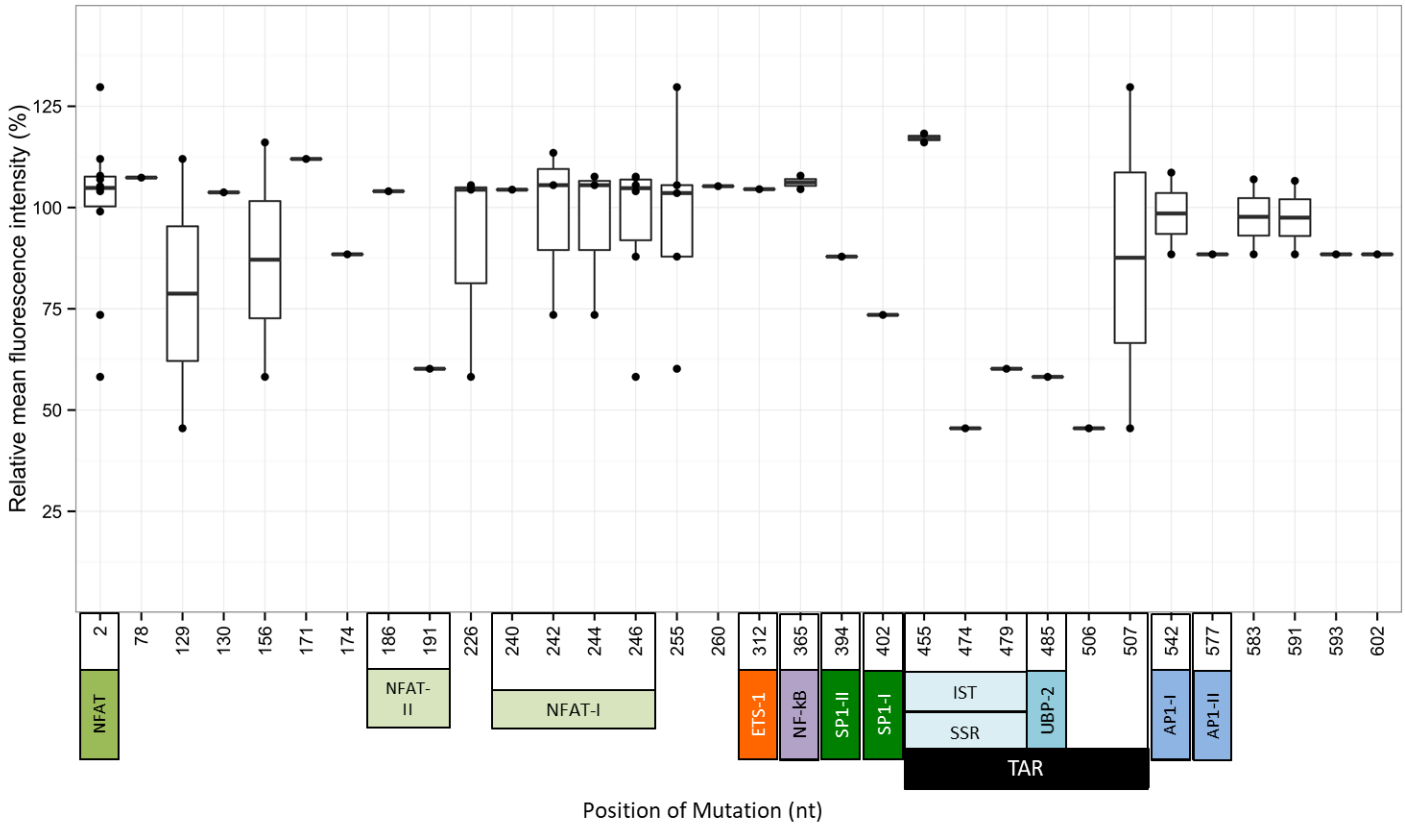

Relative mean fluorescence intensity (%)
125
100
75
50
25
NFAT
AP1 -II
COUP
NFAT-II
NFAT-I
USF-1
ETS-1
ATF-CREB
NF-kB
SP1-III
SP1 -II
SP1-I
IST
SSR
TAR
UBP-2
CTF/NFI
AP1-I
AP1-II
Position of Mutation (nt)

Figure 4.10. Box plot depicting the individual mean fluorescence intensities of the HIV-1 3' LTR A3F clones at each mutated position. The mean eGFP fluorescence intensity obtained for individual clones (dots) was plotted for each position mutated. The box plot represents the mean fluorescence intensity obtained for all clones with the specified site mutated. The location of affected transcription factor binding sites is also included.

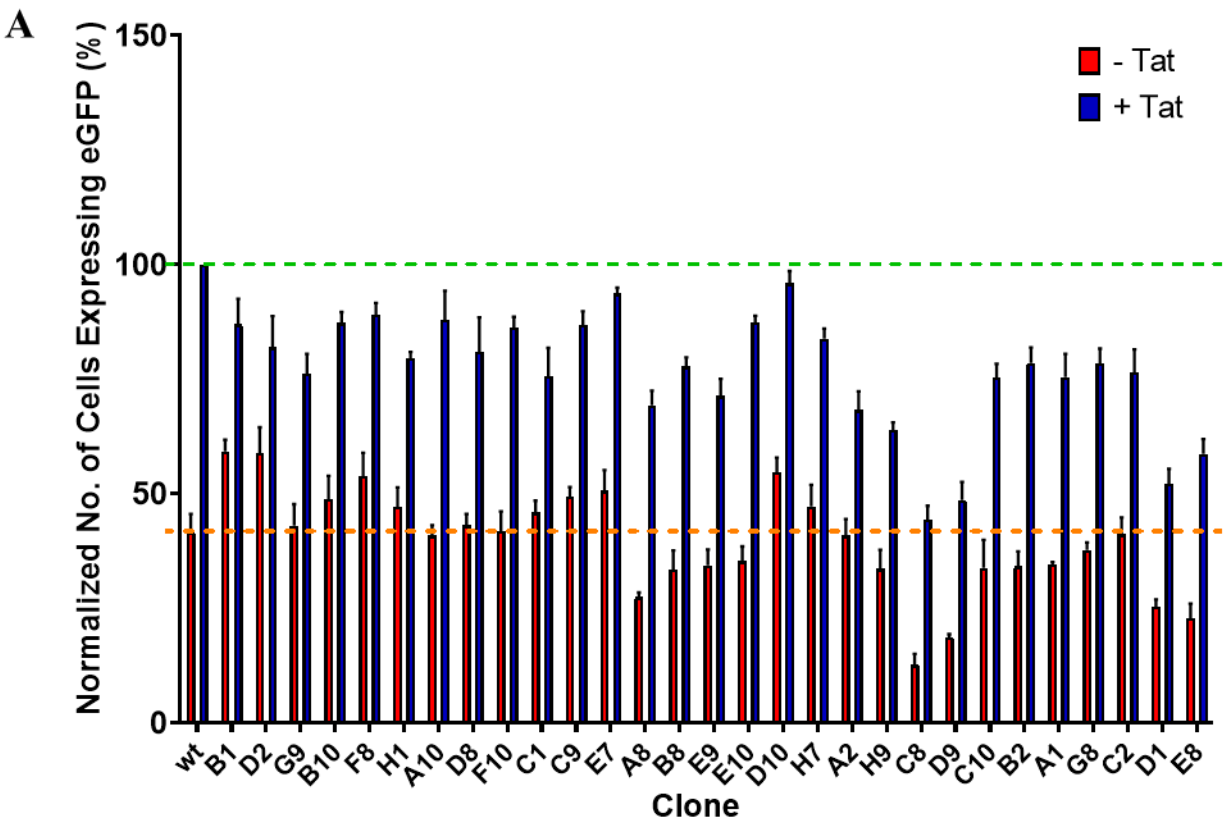

Relative mean fluorescence intensity (%)
125
100
75
50
25
2
78
129
130
156
171
174
186
191
226
240
242
244
246
255
260
312
365
394
402
455
474
479
485
506
507
542
577
583
591
593
602
NFAT
NFAT-II
NFAT-I
ETS-1
NF-kB
SP1-II
SP1-I
IST
SSR
UBP-2
TAR
AP1-I
AP1-II
Position of Mutation (nt)

4.4 Mutations in the R and U5 Regions Also Have a Modulatory Effect

Several candidate positions were selected for site-directed mutagenesis and reconstitution into HIV-1 (pNL4-3-deltaE-EGFPΔVif) ('Reconstituted Single/Double Point Mutants'). A heavily mutated reporter clone (D9) that demonstrated low eGFP fluorescence with few number of mutations was also selected for decomposition and reconstitution, in which nine out of its 17 G-to-A mutations was individually reverted via SDM to a wildtype guanine residue ('Reconstituted Revertant Clones'). These 9 positions were selected as they were in common with Clone C8, another heavily mutated clone that displayed low eGFP fluorescence with the same number of mutations (however Clone D9 was ultimately chosen for decomposition). The goal was to identify the minimum set of mutations necessary for generating a latent phenotype. The 3' LTRs of the heavily mutated reporter clones were also reconstituted into HIV-1 ('Reconstituted Heavily Mutated Clones'). A pilot infection assay was performed using a few clones selected from each set, yielding poor results (data not shown). To summarize, the revertant and heavily mutated clones were non-infectious and only the single/double point mutant clones were able to infect. We postulated that perhaps it was due to the necessity of the R region being intact in order to bind complimentarily during reverse transcription. Since the revertant and heavily mutated clones are distant from the wildtype and possess mutated R regions, this would prevent successful reverse transcription and lead to the lack of infection. This is in contrast to the single/double point mutant clones which are very close to the wildtype and generally possess intact R regions. This led us to look at the effect of mutations on the individual LTR regions separately. As mentioned previously, the U3 region houses the majority of the transcription factor binding sites. In order to determine the effects of mutations in the R and U5 regions, we cloned the mutated U3 regions of the heavily mutated clones into the eGFP expression vector, while maintaining the R and U5 regions

unmutated (Figures 4.11 and 4.13). As a brief reminder of our eGFP reporter system, the clones were transfected in the presence and absence of Tat and the levels of eGFP fluorescence were measured. Upon comparison with the fully-mutated 3' LTR clones, we noticed a difference in the MFIs between some of the latter and their partially-mutated counterparts (Figures 4.12 and 4.14). We were able to identify clones in which mutations in the R and U5 regions had a compensatory effect on promoter activity (as seen by a *decrease* in MFI of these clones versus the fully-mutated clones). Examples from the A3G clone set include clones G9, H1, A8, and H9 (which had ~0.5 or more fold decrease in MFI). Mutations in the R and U5 regions of the A3F clone set did not have much of a compensatory effect on promoter activity as there were no significant decreases in MFI of these clones versus the fully-mutated ones. We were also able to identify clones in which mutations in the R and U5 regions had a more detrimental effect on promoter activity (as seen by an *increase* in MFI of these clones versus the fully-mutated clones). Examples from the A3G clone set include clones C9, H7, D10, and G8 (which had ~1.5 or more fold increase in MFI). Examples from the A3F clone set include clones H5, F4, and D3. We did not further pursue clonings with wildtype U3-U5 nor U3-R regions with mutated R and U5 regions, respectively.

Figure 4.11. Expression assay with HIV-1 U3 mut in R-U5 wt peGFP-C3 A3G clones. The mutated U3 regions from the HIV-1 A3G clones were cloned into an eGFP expression plasmid with wildtype R-U5 regions. The effect on promoter functionality was analyzed using flow cytometry to measure the percentage of cells expressing eGFP (*A*) as well as the mean eGFP fluorescence intensity (*B*). Four independent transfections were performed and the results are presented as mean + standard error of the mean (SEM).


A
Normalized No. of Cells Expressing eGFP (%)
150
100
50
0
- Tat
+ Tat
wt B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 A8 B8 E9 E10 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone

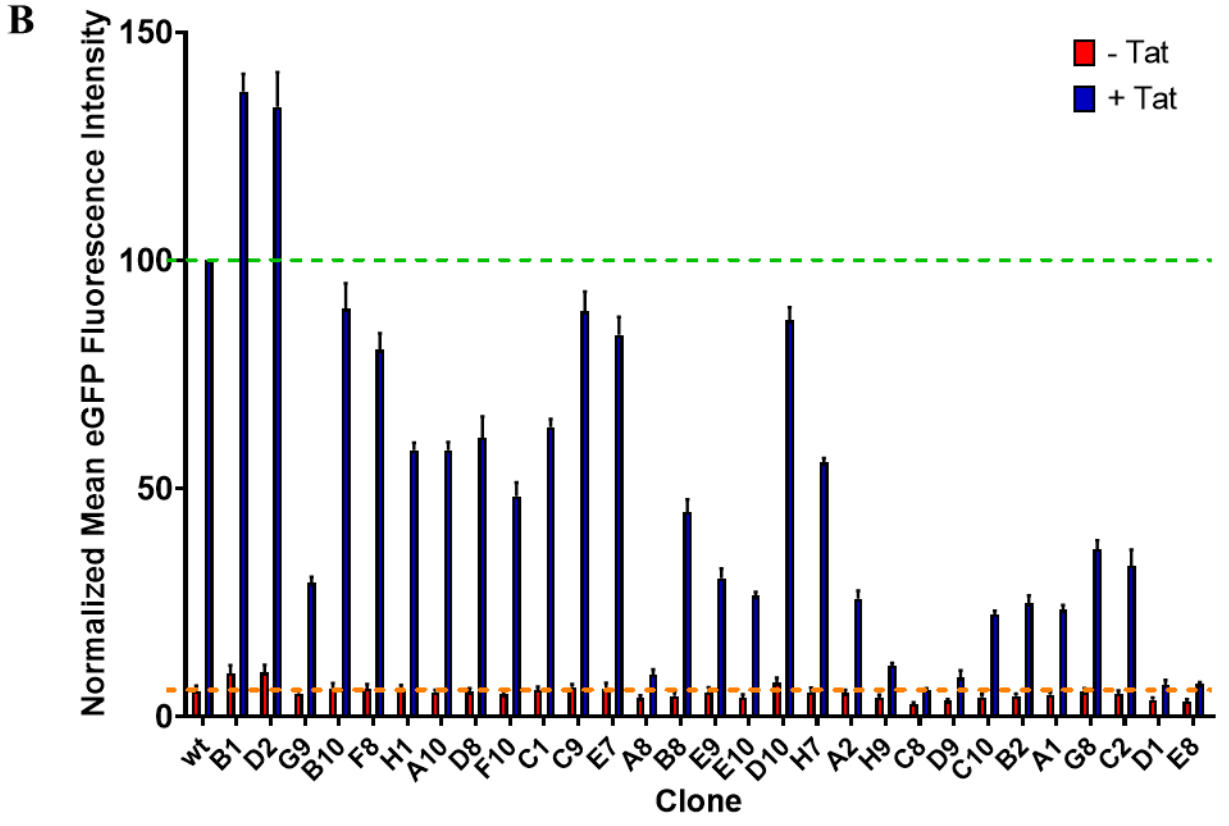
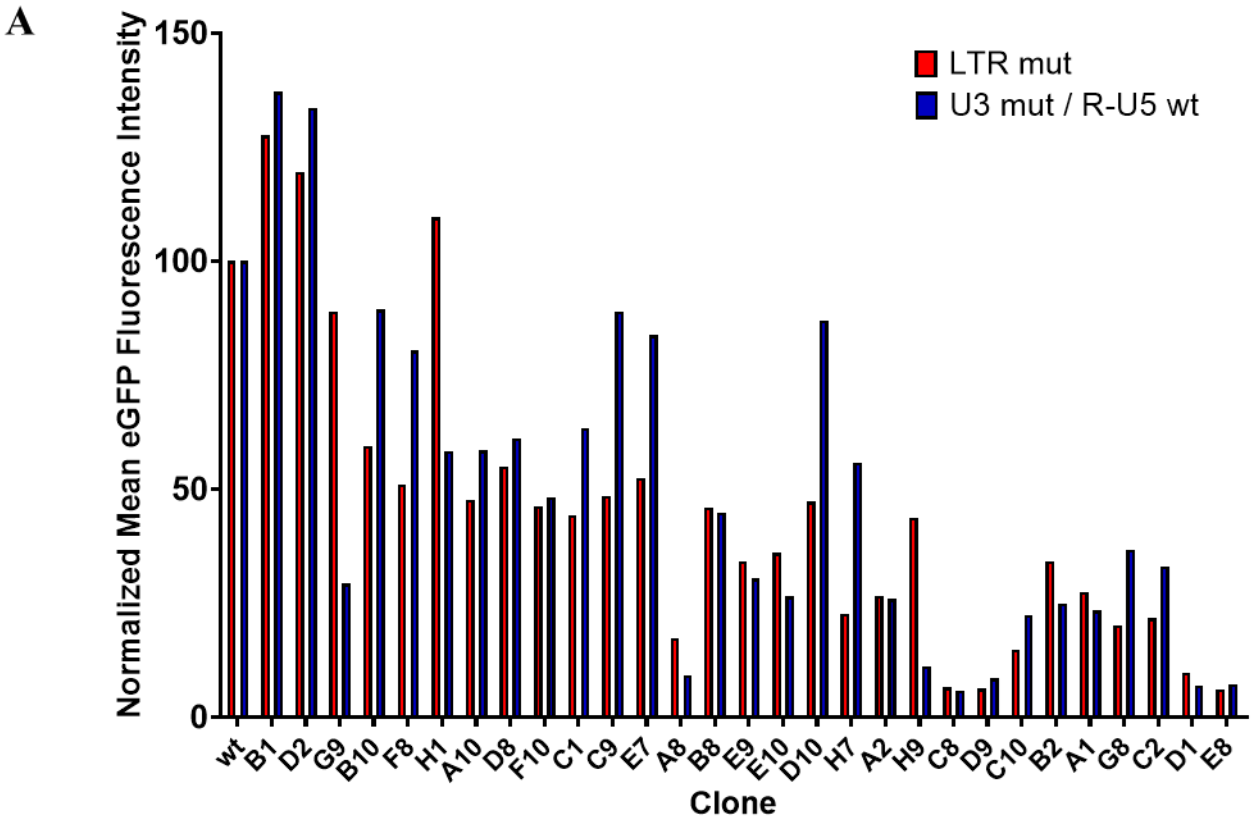

B
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
- Tat
+ Tat
wt B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 A8 B8 E9 E10 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone

Figure 4.12. Comparison of the induced mean eGFP fluorescence intensities between the HIV-1 U3 mut / R-U5 wt A3G clones and the fully-mutated 3' LTR-eGFP HIV-1 A3G clones. The Tat-induced mean eGFP fluorescence intensity of the U3 mut / R-U5 wt HIV-1 A3G clones was compared to that of the original 3' LTR-eGFP HIV-1 A3G reporter constructs (*A*). The fold difference in mean fluorescence intensity of the U3 mut / R-U5 wt clones in comparison to the fully-mutated LTR clones is also depicted (*B*). The blue dashed line indicates clearly the 1-fold mark.


A
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
LTR mut
U3 mut / R-U5 wt
wt B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 A8 B8 E9 E10 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone

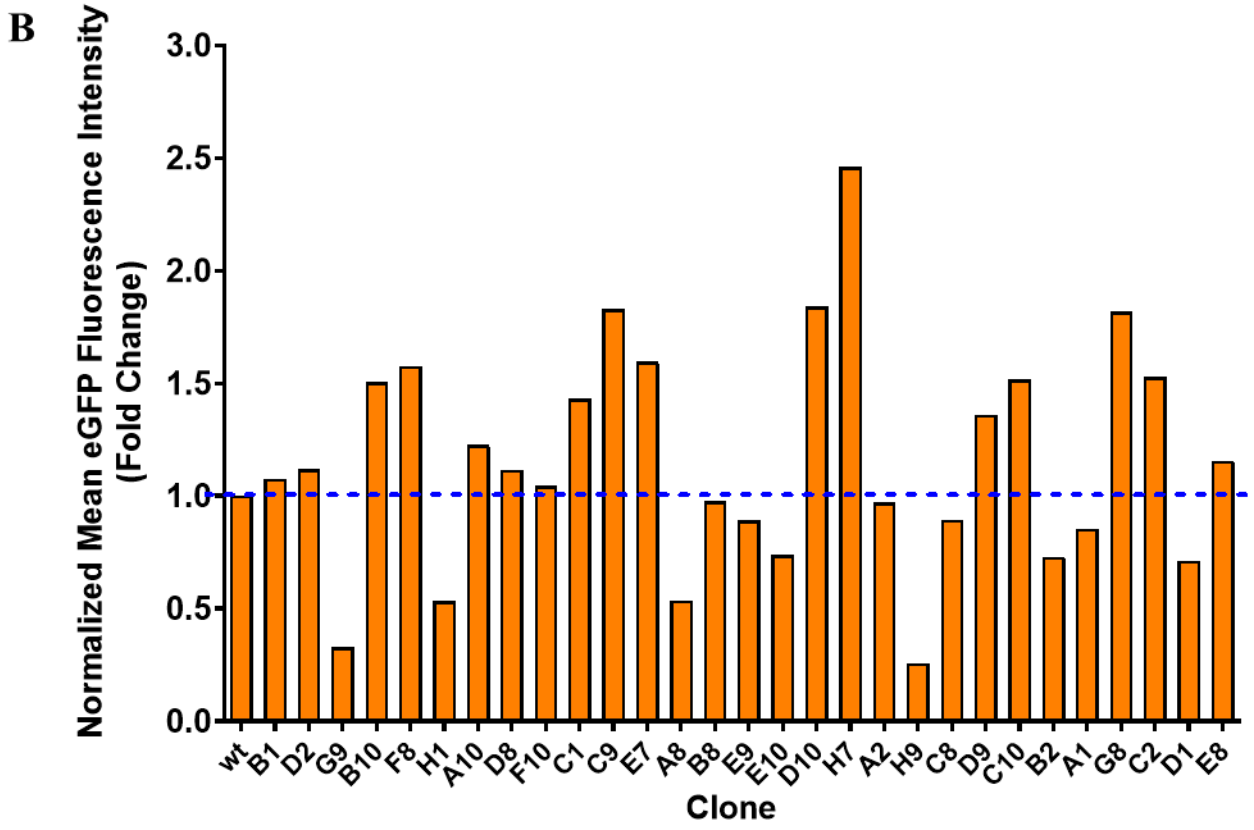
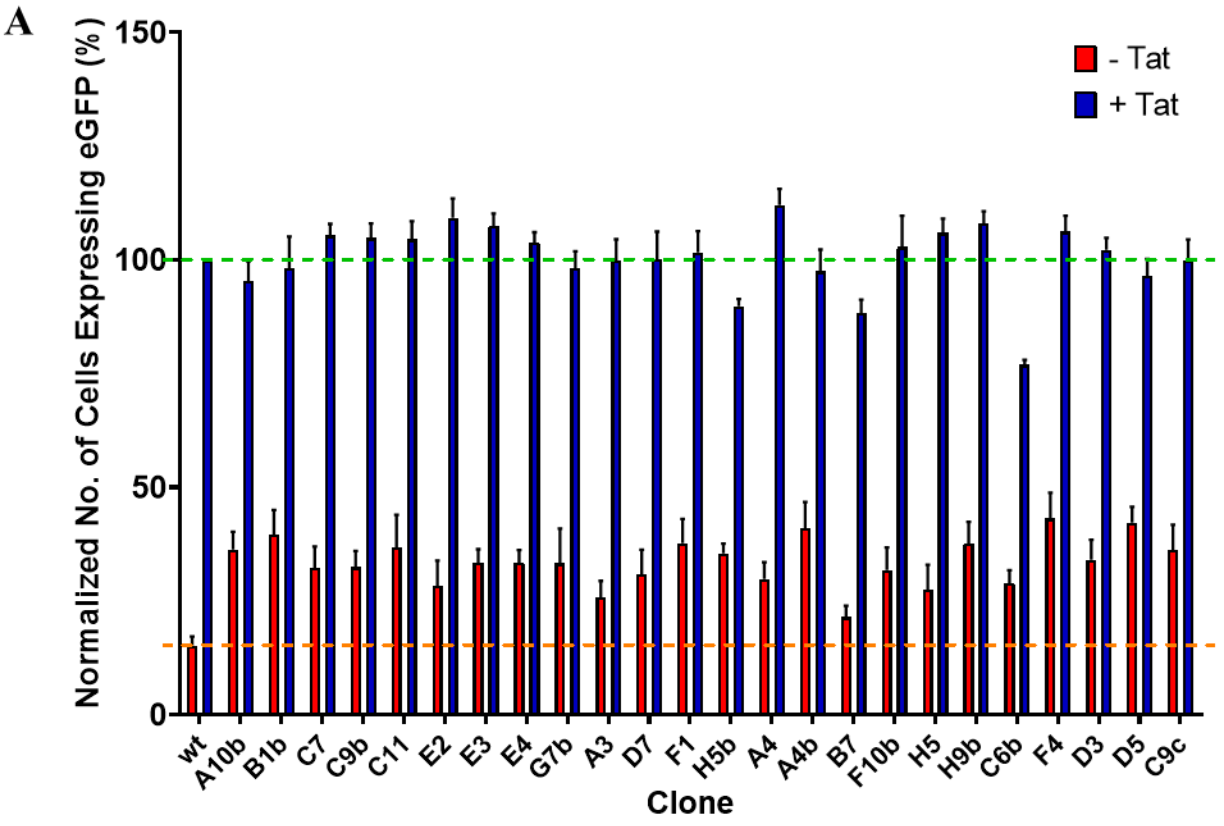

B
Normalized Mean eGFP Fluorescence Intensity
(Fold Change)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
wt B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 A8 B8 E9 E10 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone

Figure 4.13. Expression assay with HIV-1 U3 mut in R-U5 wt peGFP-C3 A3F clones. The mutated U3 regions from the HIV-1 A3F clones were cloned into an eGFP expression plasmid with wildtype R-U5 regions. The effect on promoter functionality was analyzed using flow cytometry to measure the percentage of cells expressing eGFP (*A*) as well as the mean eGFP fluorescence intensity (*B*). Three independent transfections were performed and the results are presented as mean + standard error of the mean (SEM).


A
Normalized No. of Cells Expressing eGFP (%)
150
100
50
0
- Tat
+ Tat
wt
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone

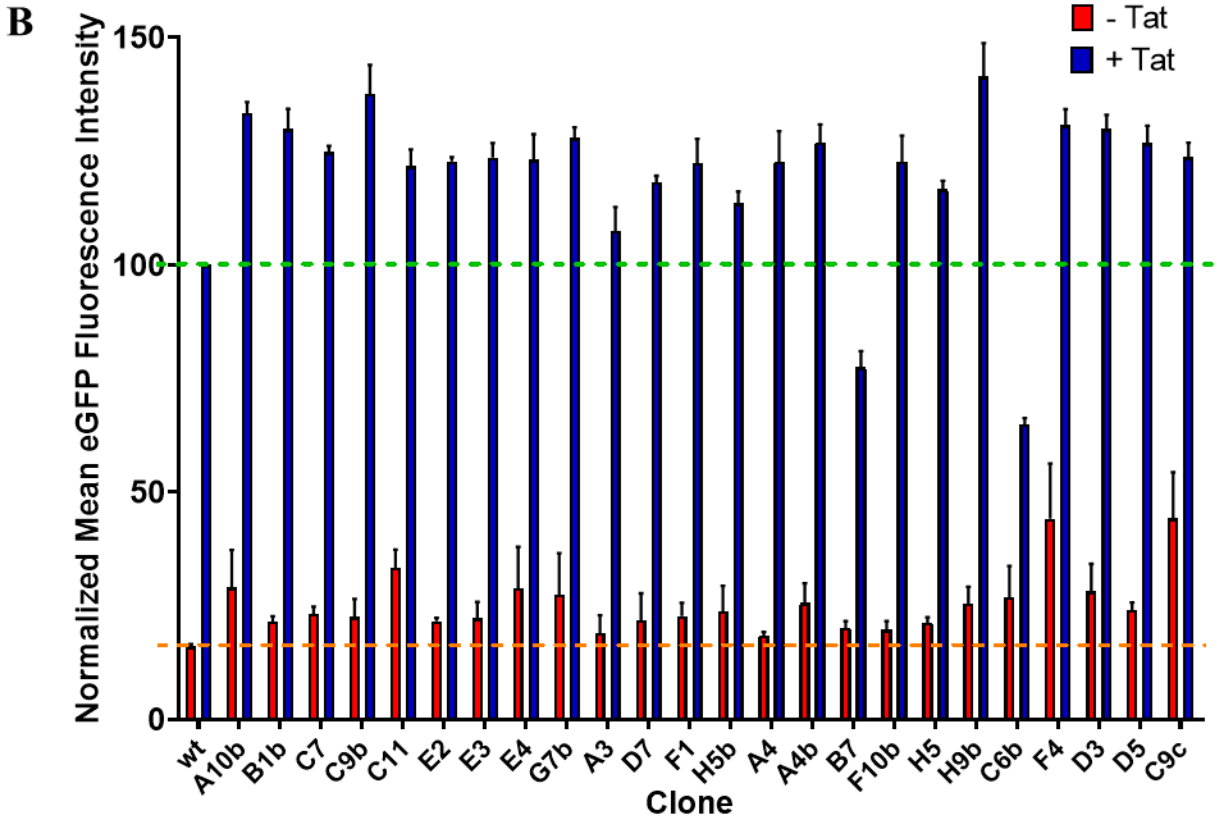
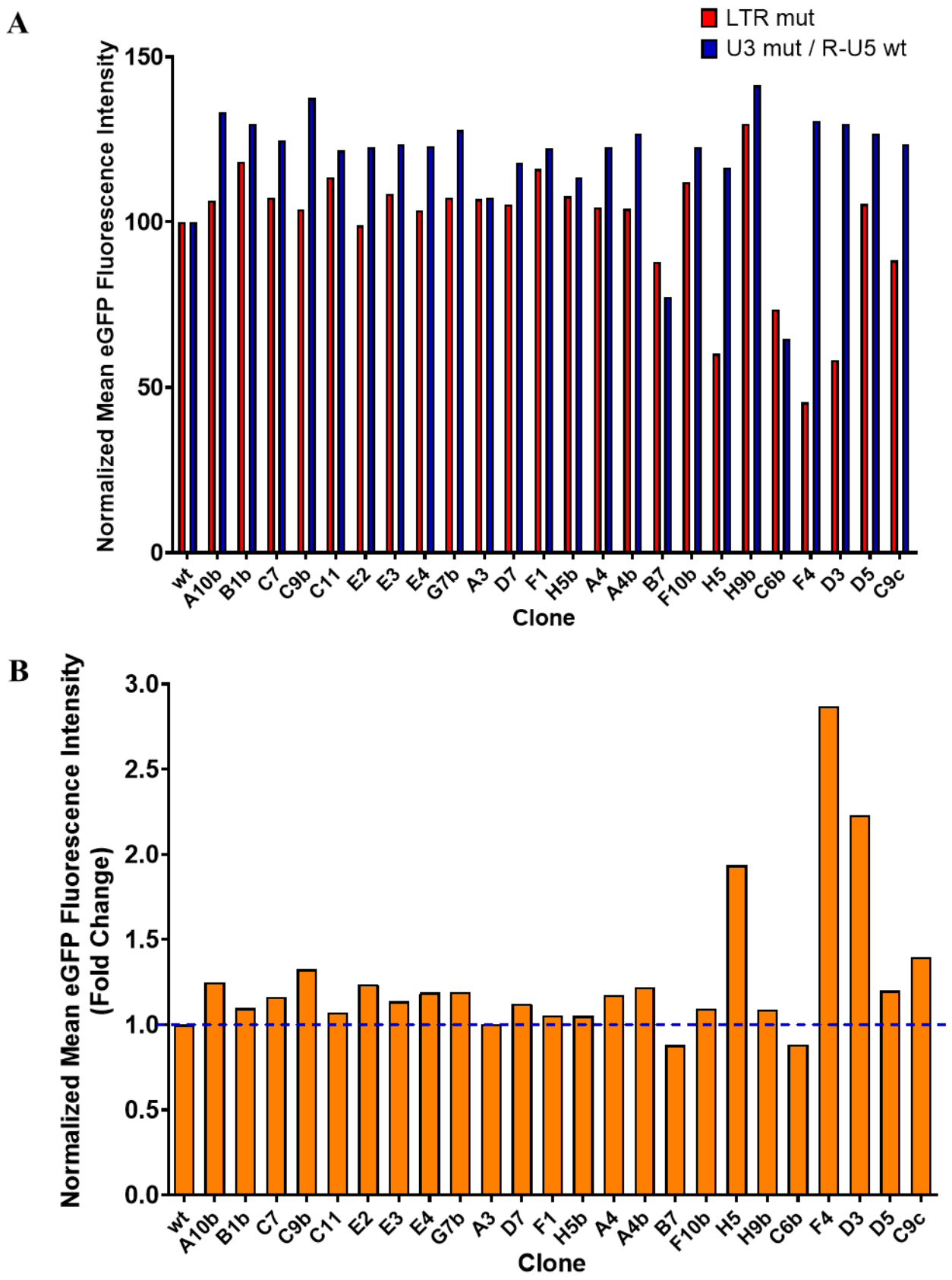

B
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
- Tat
+ Tat
wt
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone

Figure 4.14. Comparison of the induced mean eGFP fluorescence intensities between the HIV-1 U3 mut / R-U5 wt A3F clones and the fully-mutated 3' LTR-eGFP HIV-1 A3F clones. The Tat-induced mean eGFP fluorescence intensity of the U3 mut / R-U5 wt HIV-1 A3F clones was compared to that of the original 3' LTR-eGFP HIV-1 A3F reporter constructs (*A*). The fold difference in mean fluorescence intensity of the U3 mut / R-U5 wt clones in comparison to the fully-mutated LTR clones is also depicted (*B*). The blue dashed line indicates clearly the 1-fold mark.


A
LTR mut
U3 mut / R-U5 wt
Normalized Mean eGFP Fluorescence Intensity
150
100
50
0
wt
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone
B
Normalized Mean eGFP Fluorescence Intensity (Fold Change)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
wt
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone

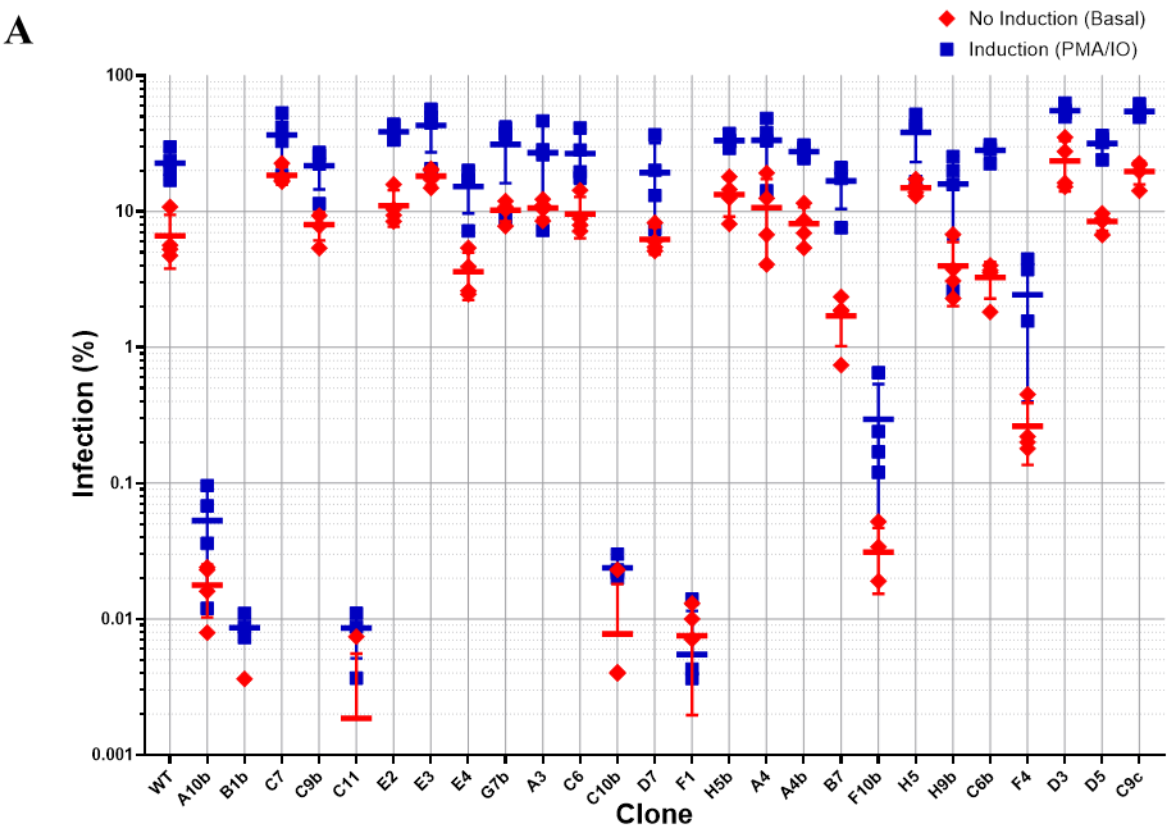
4.5 Uncovering Potentially Latent Clones

We wanted to determine whether the clones were able to be reactivated using latency reversing agents (LRAs). The initially seeming non-infectious heavily mutated clones and revertant clones may have been below the threshold initially but could perhaps be reactivated. Therefore, upon successful reconstitution of the clones, an infection assay was performed with all of the clones and their reactivation potential using phorbol myristate acetate (PMA) and ionomycin was assessed. PMA and ionomycin are routinely used in research for their synergistic effect in activating T cells (186).

When looking at the heavily mutated clones, we can see that there are in fact clones in this set that are able to express even without stimulation (Figures 4.15 and 4.16). The wildtype virus in this set had a basal infectivity of on average 6.6%. Upon PMA/ionomycin induction, the infectivity of the wildtype virus jumped to on average 22.7%, which represented a 3.4-fold increase. Having presented the clones in an order from least number of mutations to most number of mutations, we can see that there is a general trend of lower infectivity among the clones with more mutations. This is most evident in the A3G set of clones culminating in clone E8 with 37 mutations leading to the complete inactivation of the virus. However, there are outlier clones that have few number of mutations but are unable to induce (ex. clones A10b, B1b, C11, C10b, F1 in the A3F clone set and clone F10 in the A3G clone set) as well as those with a greater number of mutations but still infect well (ex. clones D3, D5, C9c in the A3F clone set and clone D10 in the A3G clone set). The heavily mutated clones with a basal infection of less than 1% were replotted to further resolve potentially latent clones (Figure 4.17). Our arbitrary parameters for identifying a latency-prone virus (LPV) were clones that had a basal infection of less than 0.25% but were able to be induced well above this threshold (~10-fold or higher). Given that we analyze on average 20 000 cells per sample during flow

acquisition, an infection of 0.25% would represent at least 50 cells being infected in the sample to be considered relevant, while also negating the chances of carryover from the previous sample during acquisition. Based on these parameters, we identified 10 clones from the heavily mutated clone set that displayed a latency-prone phenotype. These were clones A10, C1, H8, B8, E9, E10, H7, B2, C2, and H9. The latter clone displayed an impressive 220-fold increase from 0.016% (~3 cells) basal to 3.373% (~675 cells) induced infection. All ten clones came from viruses mutated by A3G. The viruses mutated by A3F were either highly infectious or inactivated, however, clones F10b and F4 just marginally missed our set parameters for a latency-prone virus and may also be of interest. A3F induced fewer mutations in comparison to A3G, however, some of the mutations were able to greatly increase the virus' infectivity, even at the basal level, compared to the wildtype control (ie. clones C7, E3, H5, D3, C9c). The mean fluorescence intensities of the heavily mutated clones tracked well with the infectivities observed. The wildtype virus had a basal MFI of on average 214 and an induced MFI of on average 349, representing a 1.6-fold increase. In general, the mean fluorescence intensities of the heavily mutated clones with a basal infection of less than 1% (and thus the LPVs) were quite low and did not increase substantially after stimulation.

Figure 4.15. Stimulation assay with reconstituted A3F heavily mutated clones. The LTRs of the A3F-mutated 3'LTR-eGFP reporter constructs were reconstituted into pNL4-3-deltaE-EGFPΔVif. HEK 293T cells were infected with each clone and stimulated with PMA and ionomycin or mock-stimulated 24 hours post-infection. The induction capacity of each clone was assessed by measuring the infectivity (*A*) and the mean fluorescence intensity (*B*) of each clone 48 hours post-infection. Clones are ordered from least number of mutations (1) to most number of mutations (7). Four independent experiments were performed and all data points are shown, including the mean ± standard deviation (SD). AU: arbitrary units.


A
No Induction (Basal)
Induction (PMA/IO)
Infection (%)
100
10
1
0.1
0.01
0.001
WT
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
C6
C10b
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone

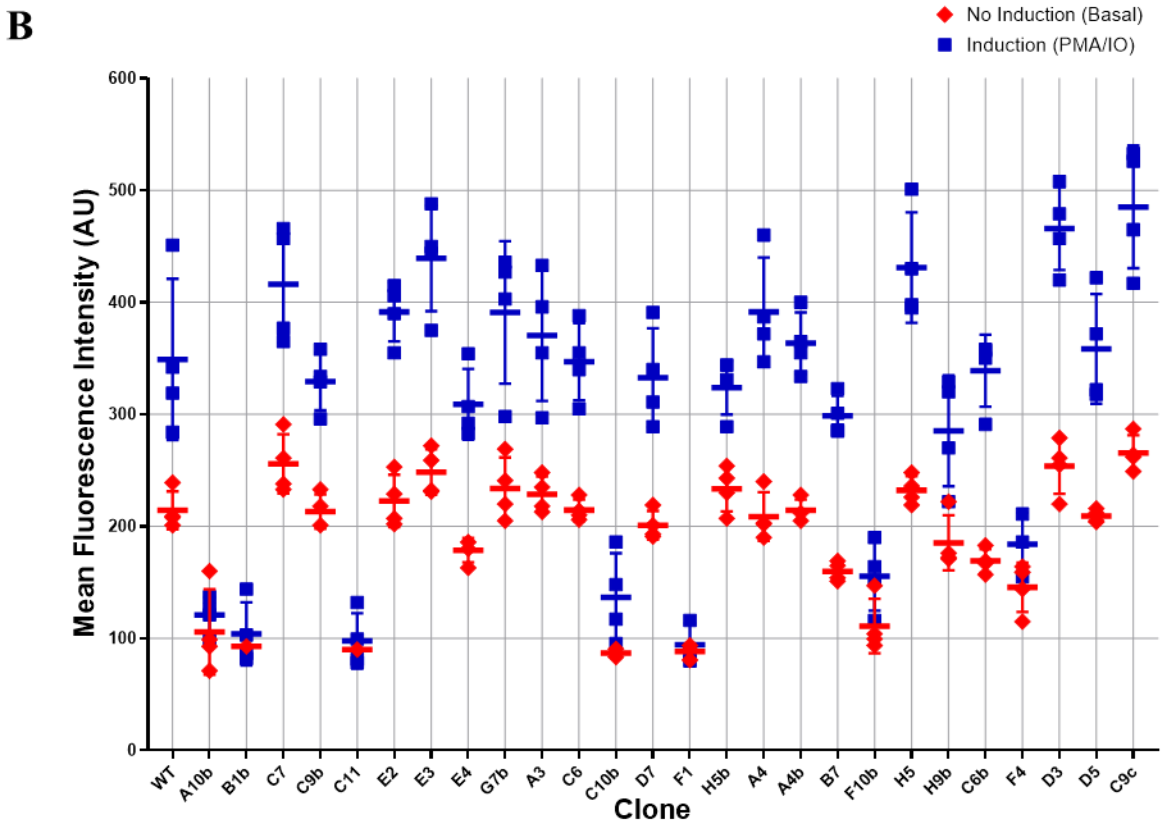
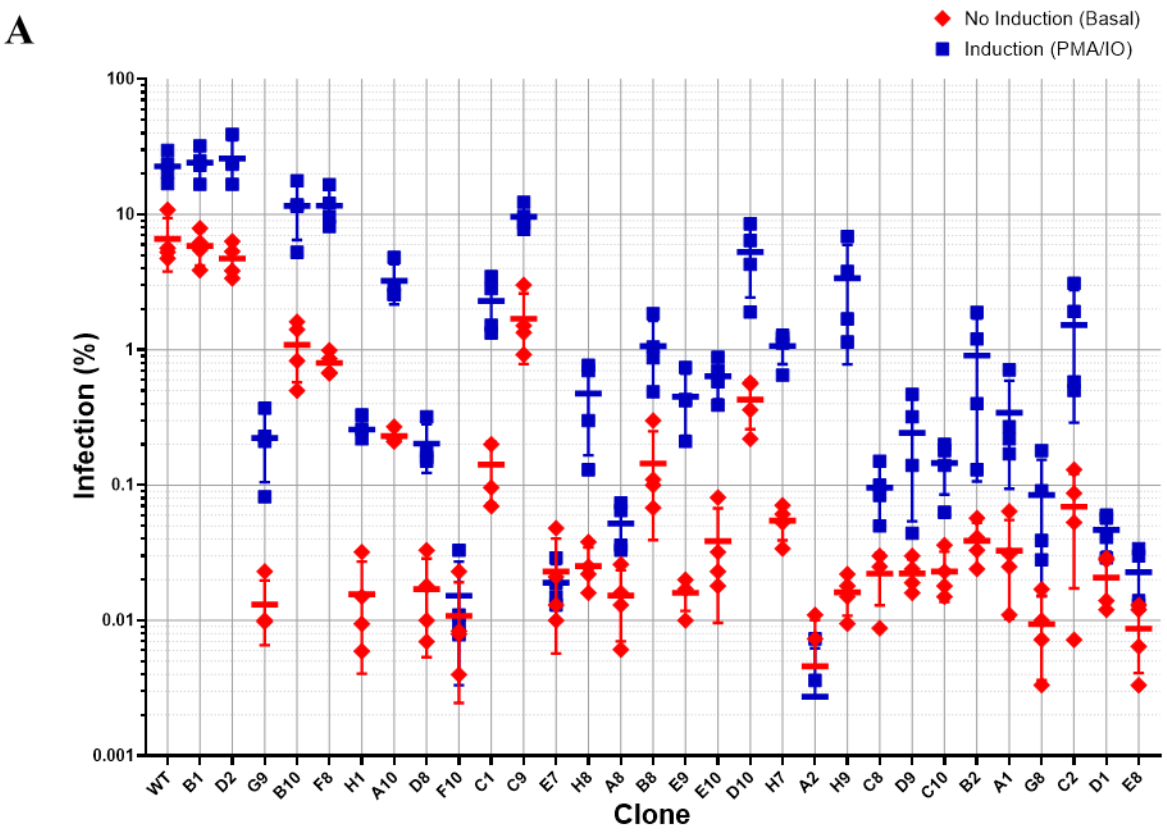

B
No Induction (Basal)
Induction (PMA/IO)
Mean Fluorescence Intensity (AU)
600
500
400
300
200
100
0
WT
A10b
B1b
C7
C9b
C11
E2
E3
E4
G7b
A3
C6
C10b
D7
F1
H5b
A4
A4b
B7
F10b
H5
H9b
C6b
F4
D3
D5
C9c
Clone

Figure 4.16. Stimulation assay with reconstituted A3G heavily mutated clones. The LTRs of the A3G-mutated 3'LTR-eGFP reporter constructs were reconstituted into pNL4-3-deltaE-EGFPΔVif. HEK 293T cells were infected with each clone and stimulated with PMA and ionomycin or mock-stimulated 24 hours post-infection. The induction capacity of each clone was assessed by measuring the infectivity (*A*) and the mean fluorescence intensity (*B*) of each clone 48 hours post-infection. Clones are ordered from least number of mutations (2) to most number of mutations (37). Four independent experiments were performed and all data points are shown, including the mean ± standard deviation (SD). AU: arbitrary units.


A
No Induction (Basal)
Induction (PMA/IO)
Infection (%)
100
10
1
0.1
0.01
0.001
WT B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 H8 A8 B8 E9 E10 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone

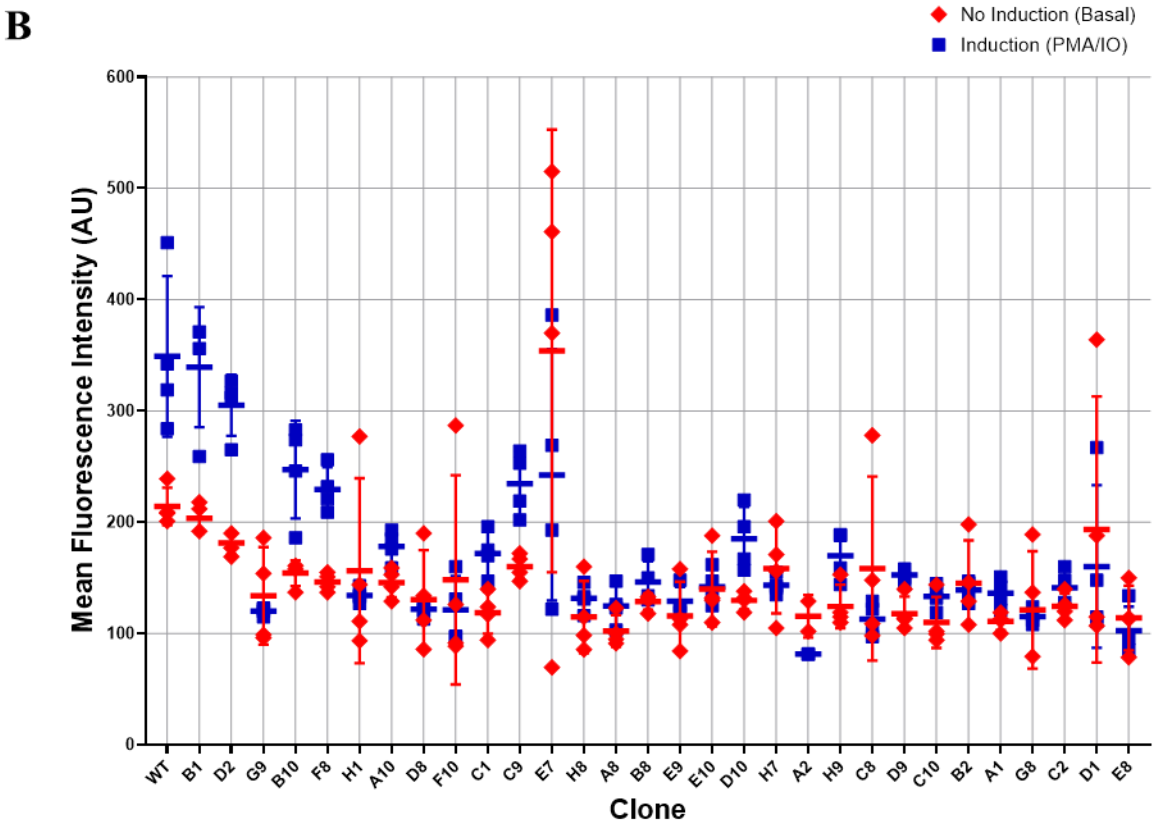
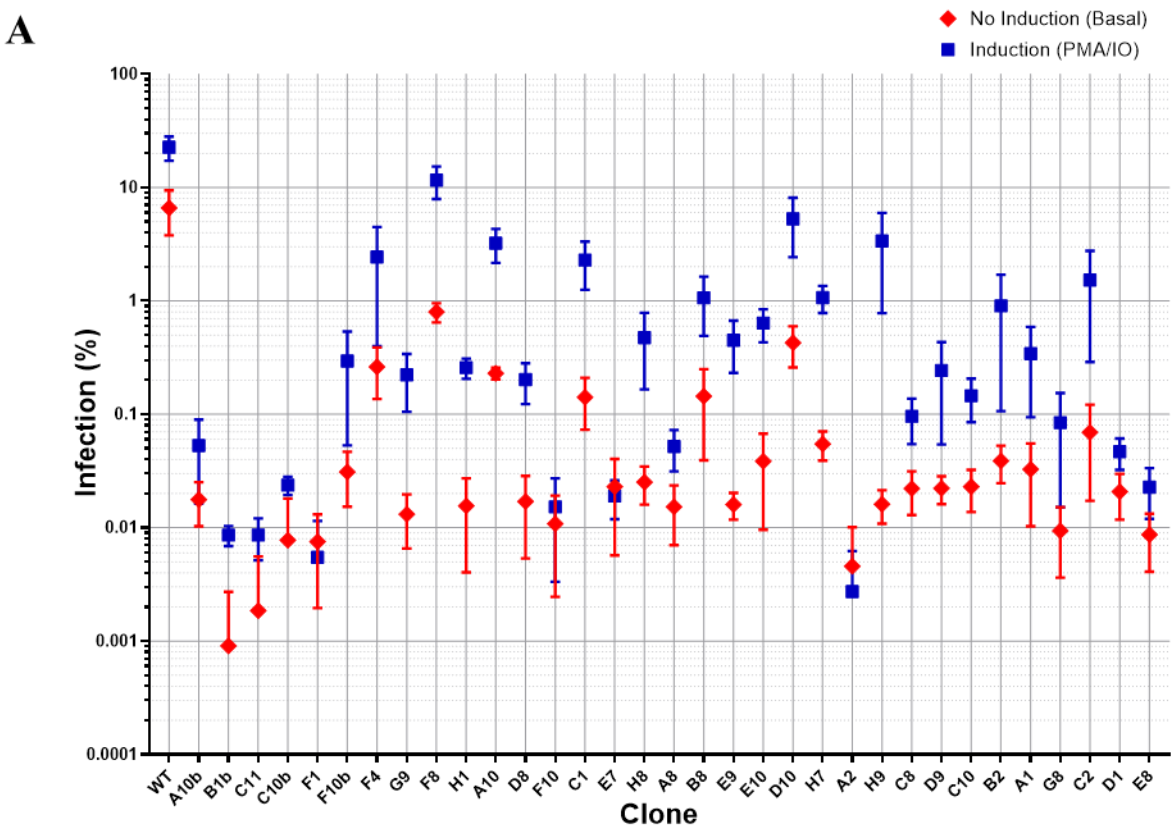
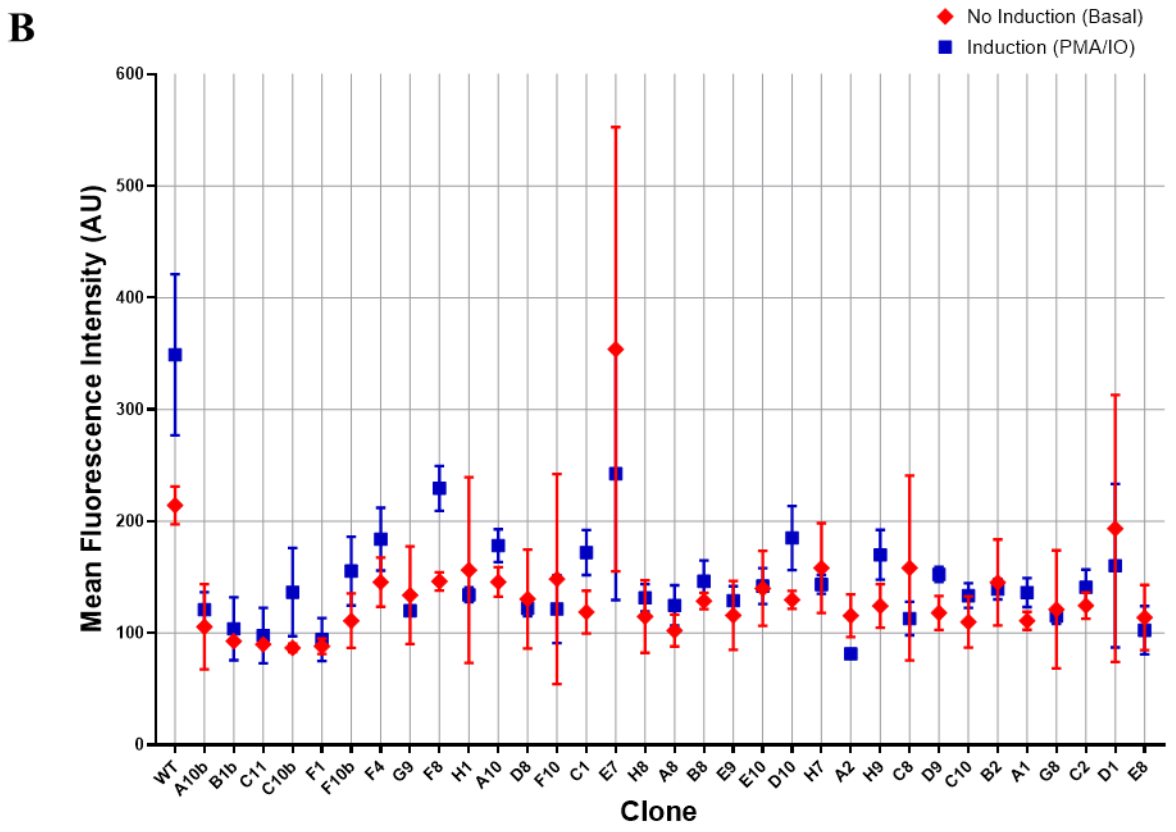

B
No Induction (Basal)
Induction (PMA/IO)
Mean Fluorescence Intensity (AU)
600
500
400
300
200
100
0
WT B1 D2 G9 B10 F8 H1 A10 D8 F10 C1 C9 E7 H8 A8 B8 E9 E10 D10 H7 A2 H9 C8 D9 C10 B2 A1 G8 C2 D1 E8
Clone

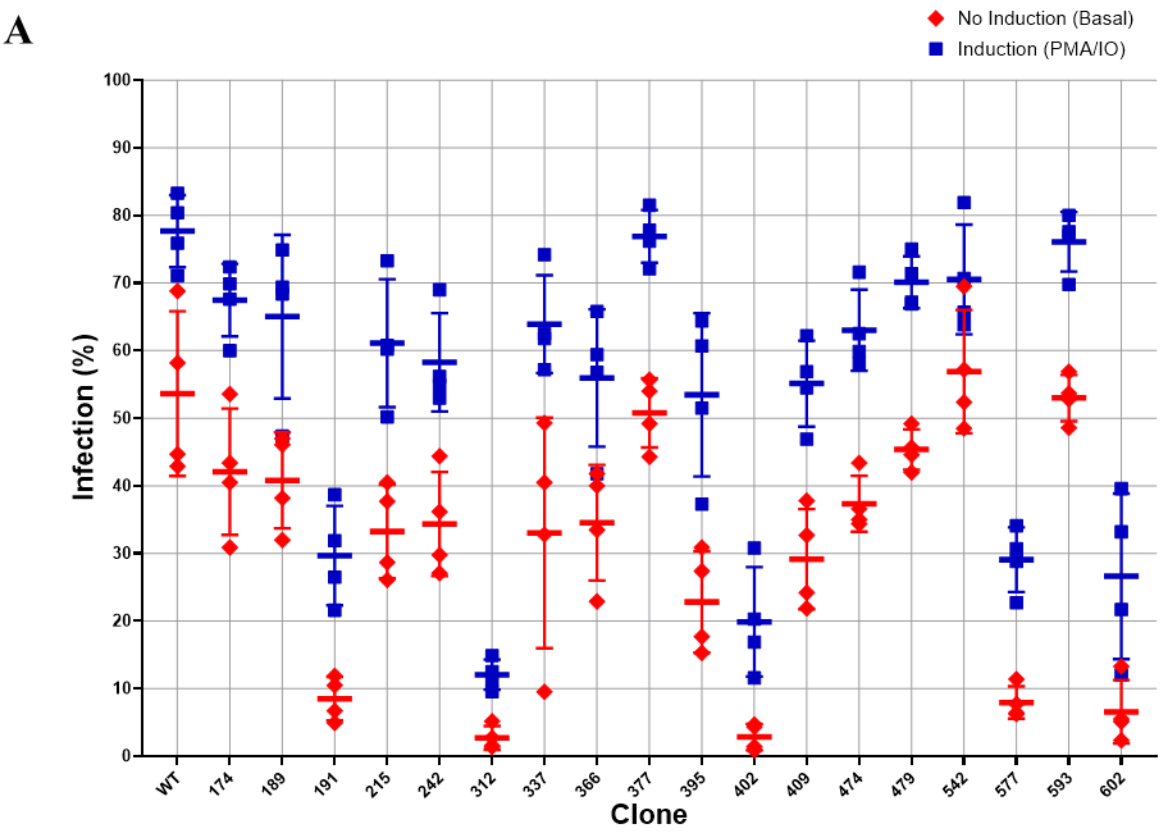
Figure 4.17. Stimulation assay with reconstituted heavily mutated clones exhibiting <1% eGFP+ cells. All clones from both the A3F and A3G sets of reconstituted heavily mutated clones that exhibited less than 1% eGFP-positive cells were replotted for clarity. The infectivity (*A*) and the mean fluorescence intensity (*B*) of each clone with and without PMA/ionomycin stimulation are displayed. Clones are ordered from least number of mutations (1) to most number of mutations (37). Four independent experiments were performed and the results are presented as mean ± standard deviation (SD). AU: arbitrary units.


A
No Induction (Basal)
Induction (PMA/IO)
Infection (%)
100
10
1
0.1
0.01
0.001
0.0001
WT
A10b
B1b
C11
C10b
F1
F10b
F4
G9
F8
H1
A10
D8
F10
C1
E7
H8
A8
B8
E9
E10
D10
H7
A2
H9
C8
D9
C10
B2
A1
G8
C2
D1
E8
Clone


B
No Induction (Basal)
Induction (PMA/IO)
Mean Fluorescence Intensity (AU)
600
500
400
300
200
100
0
WT
A10b
B1b
C11
C10b
F1
F10b
F4
G9
F8
H1
A10
D8
F10
C1
E7
H8
A8
B8
E9
E10
D10
H7
A2
H9
C8
D9
C10
B2
A1
G8
C2
D1
E8
Clone

Based on our previous pilot assay with a few reconstituted SDM clones, we know that the SDM clones with a single- or double-point mutation in the LTR of the viruses are infectious. The stimulation assays with the reconstituted single- and double-point mutants support this finding as the trend still holds true for all clones tested (Figures 4.18 and 4.19). The wildtype control in this set of clones had a basal infectivity of on average 53.7% and an induced infectivity of on average 77.7%, representing a 1.5-fold increase. The majority of the clones were highly infectious with only single-mutant clones 191, 312, 402, 506, 577, 602, and double-mutant clones 350/454 and 393/454 displaying infection rates noticeably lower than the wildtype control (≤10% basal and ≤30% induced). Interestingly, these slightly impaired clones were also the ones most responsive to stimulation when looking at fold difference in induction (Figure 4.20). For example, the clone with a single mutation at position 402 was induced 11-fold from a basal infectivity of 2.9% to a stimulated infectivity of 19.9%. When looking at the mean fluorescence intensities of the SDM clones in comparison to the wildtype control, we can see that the MFIs track well with the infectivities observed. The wildtype control had a basal MFI of on average 454 and an induced MFI of on average 736, representing a 1.7-fold increase. The MFIs of the above-mentioned clones were significantly lower than the wildtype control, not even reaching an MFI of 300 after induction. On the other hand, we also identified clones where a single- or double-point mutation in the LTR allowed the virus to respond to stimulation at a higher fold than the wildtype control (≥1.8-fold induction, ie. clones 363, 377, 379, 406). Although the induction capacity of these clones was higher, they were still within similar levels of fluorescence as the wildtype control and did not surpass the MFI of the wildtype at neither the basal nor induced levels.

Figure 4.18. Stimulation assay with reconstituted A3F single-point mutants. The LTR of pNL4-3-deltaE-EGFPΔVif was synthetically-mutated at specific A3F deamination target sites. Clones contain one single mutation located only in the 3' LTR. HEK 293T cells were infected with each clone and stimulated with PMA and ionomycin or mock-stimulated 24 hours post-infection. The induction capacity of each clone was assessed by measuring the infectivity (*A*) and the mean fluorescence intensity (*B*) of each clone 48 hours post-infection. Four independent experiments were performed and all data points are shown, including the mean ± standard deviation (SD). AU: arbitrary units.


A
No Induction (Basal)
Induction (PMA/IO)
Infection (%)
100
90
80
70
60
50
40
30
20
10
0
WT
174
189
191
215
242
312
337
366
377
395
402
409
474
479
542
577
593
602
Clone

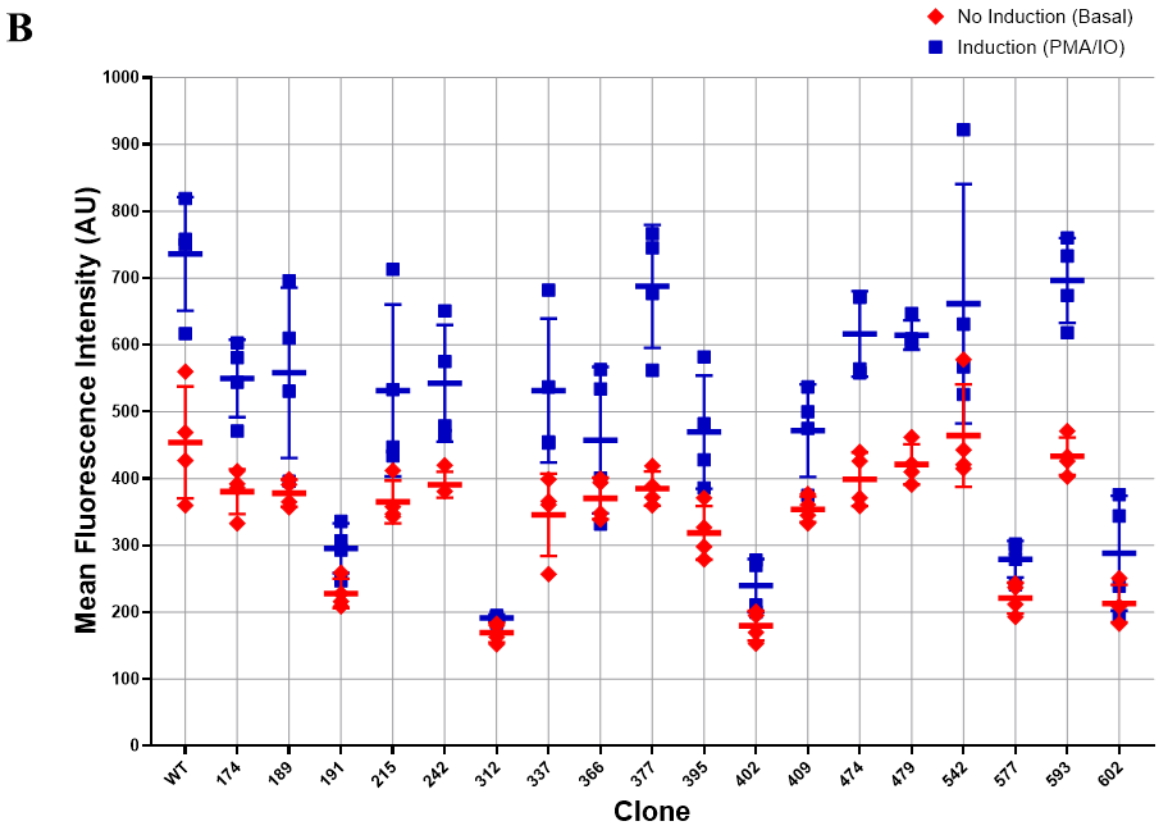
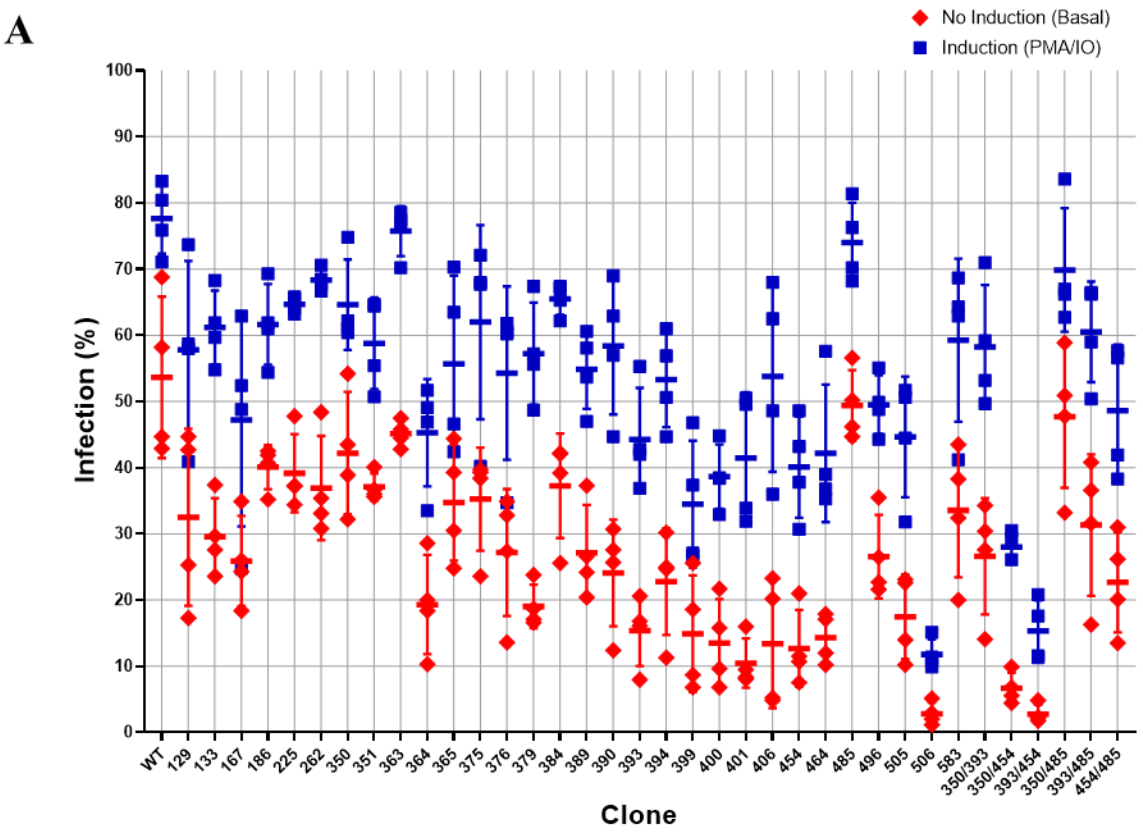

B
No Induction (Basal)
Induction (PMA/IO)
Mean Fluorescence Intensity (AU)
1000
900
800
700
600
500
400
300
200
100
0
WT
174
189
191
215
242
312
337
366
377
395
402
409
474
479
542
577
593
602
Clone

Figure 4.19. Stimulation assay with reconstituted A3G single- and double-point mutants. The LTR of pNL4-3-deltaE-EGFPΔVif was synthetically-mutated at specific A3G deamination target sites. Clones contain one or two mutations located only in the 3' LTR. HEK 293T cells were infected with each clone and stimulated with PMA and ionomycin or mock-stimulated 24 hours post-infection. The induction capacity of each clone was assessed by measuring the infectivity (*A*) and the mean fluorescence intensity (*B*) of each clone 48 hours post-infection. Four independent experiments were performed and all data points are shown, including the mean ± standard deviation (SD). AU: arbitrary units.


A
No Induction (Basal)
Induction (PMA/IO)
Infection (%)
Clone
WT
129
133
167
186
225
262
350
351
363
364
365
375
376
379
384
389
390
393
394
399
400
401
406
454
464
485
496
505
506
583
350/393
350/454
393/454
350/485
393/485
454/485

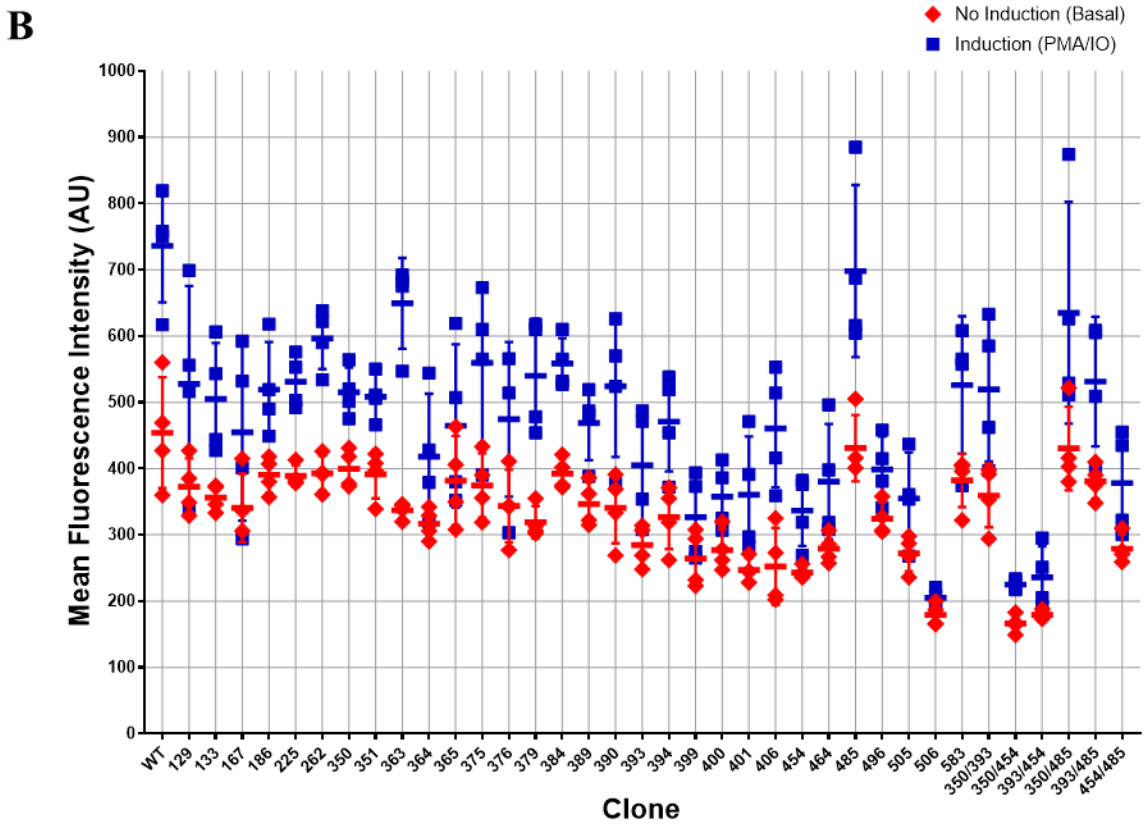
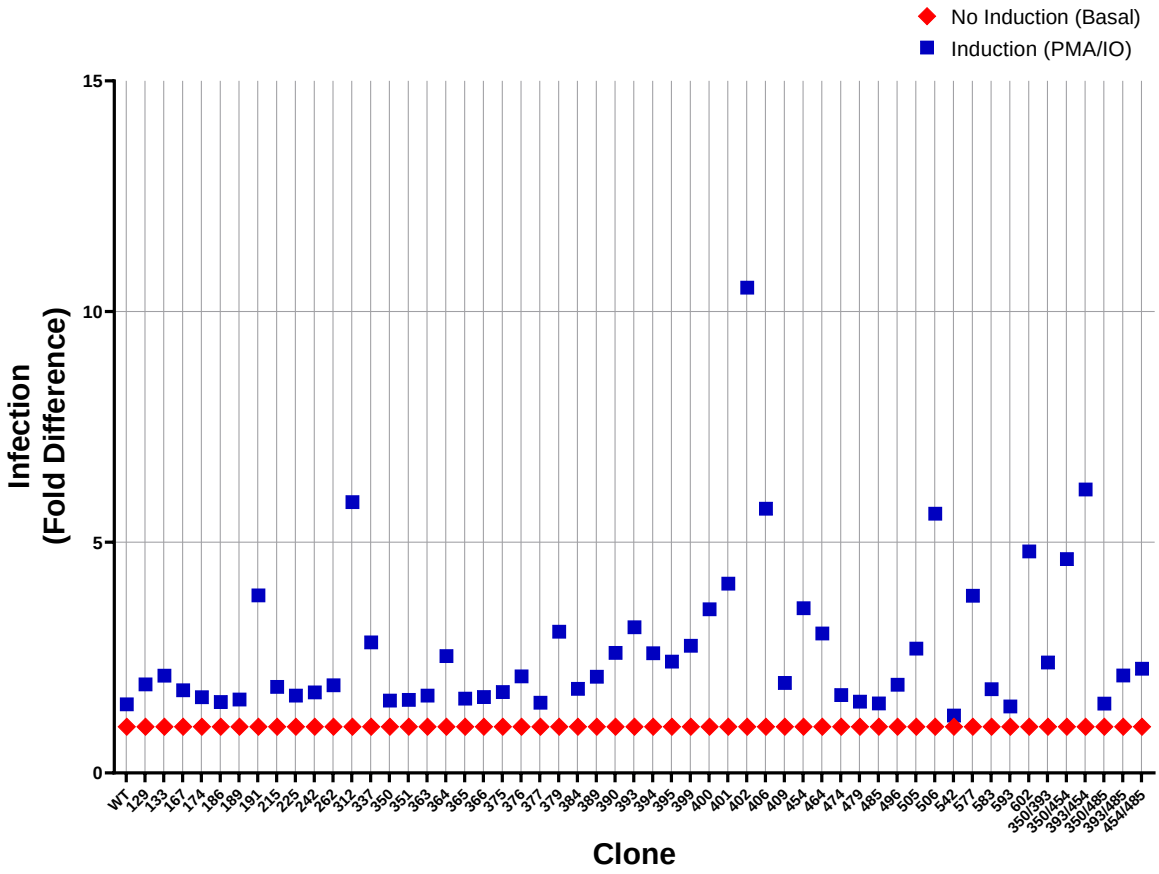
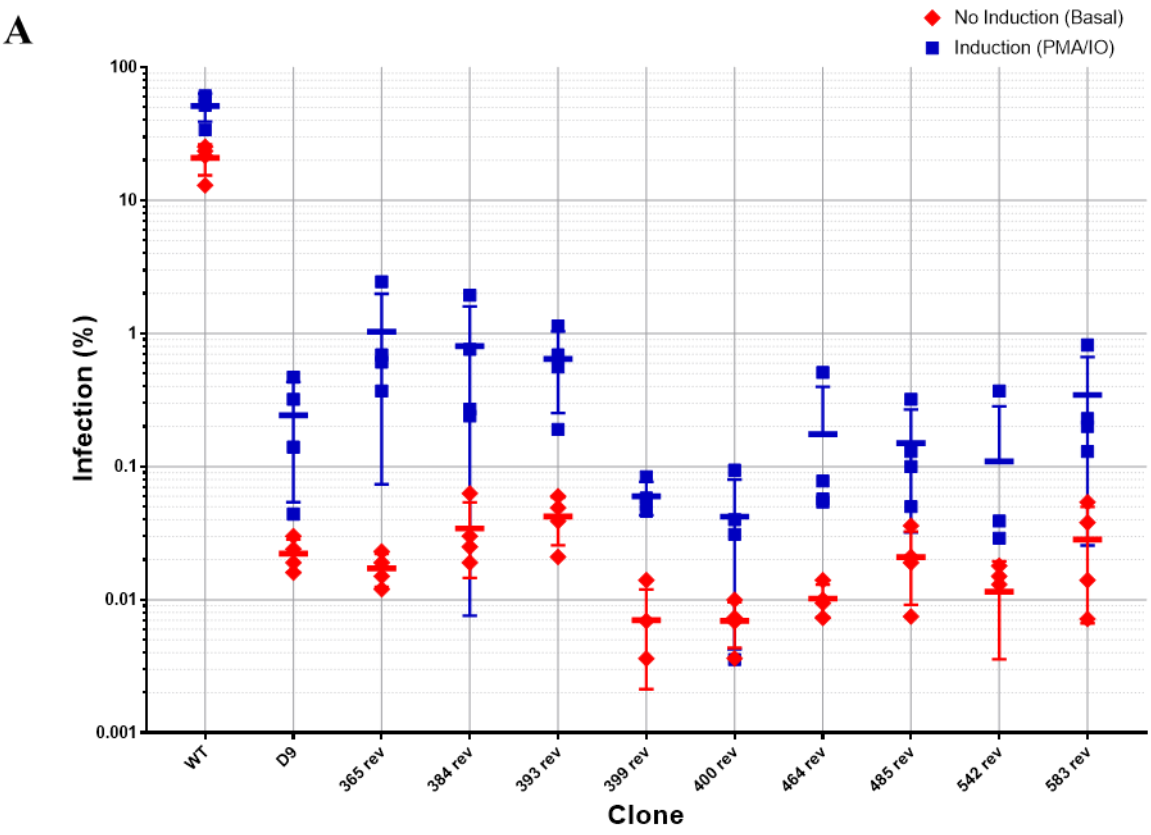

B
No Induction (Basal)
Induction (PMA/IO)
Mean Fluorescence Intensity (AU)
Clone
WT
129
133
167
186
225
262
350
351
363
364
365
375
376
379
384
389
390
393
394
399
400
401
406
454
464
485
496
505
506
583
350/393
350/454
393/454
350/485
393/485
454/485

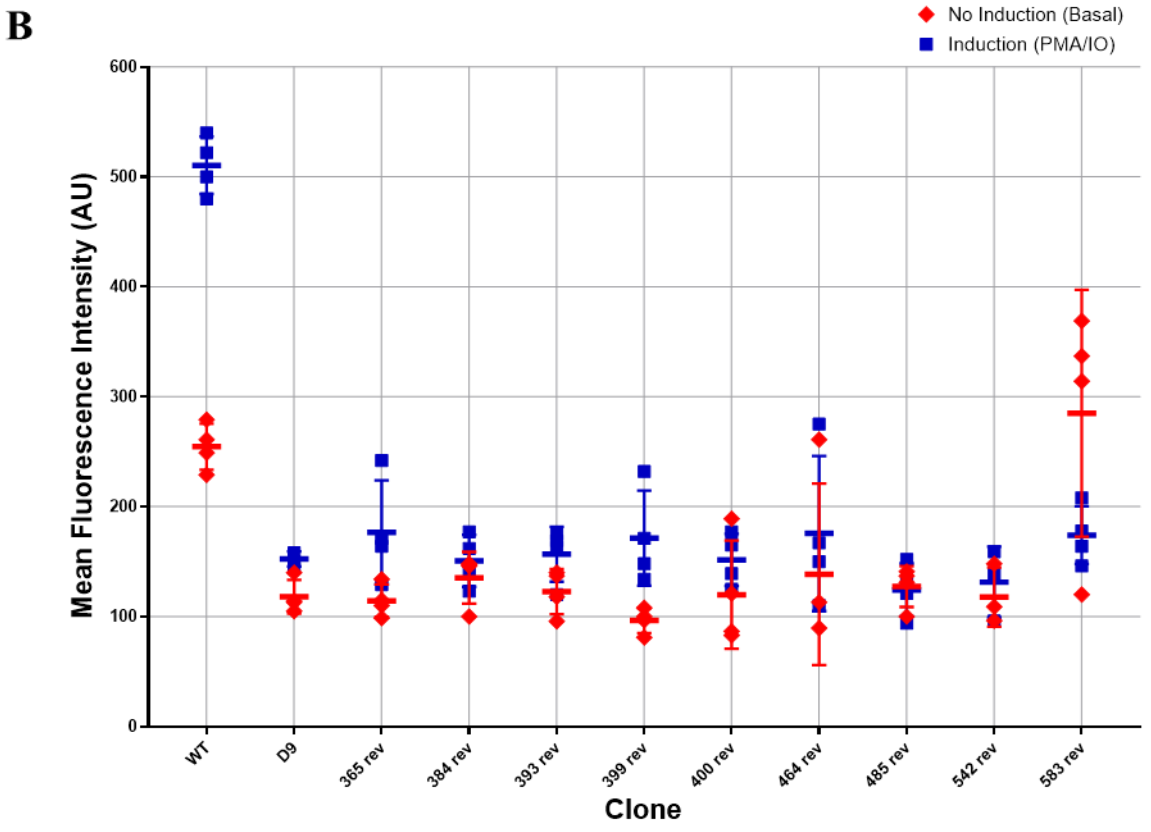
Figure 4.20. Fold induction in infectivity among all reconstituted single- and double-point mutants. The mean fold difference in infectivity from basal to stimulated of all reconstituted SDM clones is plotted.


No Induction (Basal)
Induction (PMA/IO)
Infection
(Fold Difference)
15
10
5
0
WT
129
133
167
174
186
189
191
215
225
242
262
312
337
350
351
363
364
365
366
375
376
377
379
384
389
390
393
394
395
399
400
401
402
406
409
454
464
474
479
485
496
505
506
542
577
583
593
602
350/393
350/454
393/454
350/485
393/485
454/485
Clone

As a reminder, the revertant clones were generated by individually converting nine out of 17 of the G-to-A mutations found in heavily mutated clone D9 into wildtype guanine residues. Clone D9 was chosen because it, along with clone C8, had a few number of mutations while demonstrating low eGFP fluorescence. The goal was to decompose one of these clones (clone D9 was ultimately chosen) to eventually reach a minimum set of mutations that would generate a latent phenotype. In the early pilot infection assay with a few of the revertant clones, it was apparent that these clones were non-infectious. We tested their reactivation potential using PMA and ionomycin and found that the clones were still poorly-infectious despite stimulation (Figure 4.21). The wildtype virus in this clone set had a basal infectivity of on average 20.9% and an induced infectivity of on average 51.3%, representing a 2.5-fold increase. Clone D9 had a basal infectivity of on average 0.022% and an induced infectivity of on average 0.244%, representing a 12-fold increase. None of the singly reverted mutations were able to rescue the phenotype of clone D9 towards the wildtype control, with the highest infection reaching only 1% (clones 365 rev, 384 rev, 393 rev). The reversion from A-to-G at positions 399 and 400 in clone D9 appeared to impair the clone further as the infectivities of these two clones (399 rev, 400 rev) were slightly lower than the non-decomposed clone D9. The four remaining clones (464 rev, 485 rev, 542 rev, 583 rev) displayed infectivities similar to the non-decomposed clone D9. The mean fluorescence intensities of the revertant clones were correspondingly quite low in comparison to the wildtype control. The wildtype virus had a basal MFI of on average 255 and an induced MFI of on average 511, representing a 2-fold increase. Clone D9 had a basal MFI of on average 118 and an induced MFI of on average 153, representing a 1.3-fold increase. All of the revertant clones had low MFI levels similar to the non-decomposed clone D9.

Figure 4.21. Stimulation assay with reconstituted revertant clones. Nine G-to-A mutations in the LTR of clone D9 were individually reverted to a wildtype G residue and reconstituted into pNL4-3-deltaE-EGFPΔVif. Each revertant clone has 16 A3-induced mutations and 1 reverted A-to-G mutation. HEK 293T cells were infected with each clone and stimulated with PMA and ionomycin or mock-stimulated 24 hours post-infection. The induction capacity of each clone was assessed by measuring the infectivity (*A*) and the mean fluorescence intensity (*B*) of each clone 48 hours post-infection. Four independent experiments were performed and all data points are shown, including the mean ± standard deviation (SD). AU: arbitrary units.


A
No Induction (Basal)
Induction (PMA/IO)
Infection (%)
100
10
1
0.1
0.01
0.001
WT
D9
365 rev
384 rev
393 rev
399 rev
400 rev
464 rev
485 rev
542 rev
583 rev
Clone


B
No Induction (Basal)
Induction (PMA/IO)
Mean Fluorescence Intensity (AU)
600
500
400
300
200
100
0
WT
D9
365 rev
384 rev
393 rev
399 rev
400 rev
464 rev
485 rev
542 rev
583 rev
Clone

5.0 DISCUSSION

HIV-1 has proven to be one of the most difficult pathogens humans have had to face. Despite continuous efforts, the virus is still incurable and there is currently no vaccine against it. The biggest barrier to eradicating this pathogen is the virus' ability to persist in latent reservoirs which prevent it from being detected by the host immune system and antiretroviral therapies. Add to this the possibility that restriction factors meant to protect us against the virus can also be aiding the virus in its elusive behaviour and we have ourselves one formidable pathogen.

APOBEC3 proteins mutate all along the genome of HIV-1 (174), however, my project focussed on mutations only in the LTR promoter of the virus. My first task was to demonstrate that the HIV-1 LTR was being significantly targeted by the A3 proteins. Although four A3s function against HIV-1, I worked with A3G and A3F only as they are the two most potent A3s against HIV-1 as well as possessing different deamination contexts. To aid in my investigation, I generated a library of clones that only carried A3-induced mutations in the LTR promoter of the virus. Naturally, my first efforts were to clone the 5' LTR of the viruses that were exposed to A3G and A3F since it is the 5' LTR that commonly acts as the promoter for the virus and my aim was to study the effects of A3-induced mutations on this regulatory region. Unfortunately, it became apparent that the 5' LTR is not mutated in a single round of HIV-1 infection and a second round of infection must take place. In the first round of infection, it is the 3' LTR that is mutated by the APOBEC3 proteins. In a subsequent round of infection, these mutations would be transferred over to the 5' LTR during reverse transcription and this is when the A3-induced mutations would be able to enact their effects on the viral promoter. The clones that were mutated by A3F did not produce many mutations (between 1 to 7) and had a modest effect on promoter activity when assessing the expression of the eGFP reporter

gene and its mean fluorescence intensity. However, the clones mutated by A3G were much more abundantly-mutated (possessing between 2-37 mutations) and these mutations had a wide range of effects on the promoter of the clones. From these promoter replacement assays, we were able to observe some clones that possessed mutations that not only affected their responsiveness to Tat, but also disrupted their basal transcription as well. Some of these clones were no longer inducible by Tat and were completely inactivated by the mutations they acquired. Yet in some instances, a few clones possessed mutations that allowed them to express eGFP slightly better than even the wildtype control. The clones we were interested in, however, were ones that were weakly-transcribing but were still able to be activated by Tat induction.

In order to decipher their method of attenuation, we decided to look at the location of the mutations generated. The LTR promoter is home to many transcription factor binding sites, especially in the U3 region. I generated maps of transcription factor binding sites that were placed to scale above tick plots of putative deamination target sites as well as cumulative confirmed target sites for all clones. From these maps, it is evident that the mutations cluster beneath important transcription factor binding sites for master regulators of transcription such as NF-κB, Sp1, NFAT, AP-1, and also the TAR element. When looking at hot spots for deamination, we also noticed that NF-κB and TAR were frequently mutated at more than one location in their sequence. Because these transcription factors are so crucial, as well as Tat binding to the TAR element, any mutations in the binding sites and sequences could have a direct impact on the efficiency of transcription of the viruses. Improper binding of transcription factors and the lack of Tat transactivation could lead to weak transcription and the lack of viral antigens recognized by the immune system and antiviral therapeutics. The next step in studying the impact of these mutations on transcription factor binding would be to perform chromatin

immunoprecipitation (ChIP) assays to determine the degree to which transcription factor binding is being affected. ChIP assays are a powerful tool used to probe protein-DNA interactions (187). The protein-DNA interactions are temporarily cross-linked using formaldehyde to preserve their interaction. The chromatin is then fragmented using a nuclease, and antibodies specific to the proteins are used to precipitate them. Because the proteins are complexed with the DNA, it is co-precipitated and the entire complex is captured by beads. The protein-DNA cross-linking is then reversed, and the DNA is purified. The purified DNA can then be used for downstream analyses such as its quantification. This technique would be useful in measuring the impact the A3-induced mutations in the binding sites have on the ability of the transcription factors to bind to their DNA target sequences. Similarly, the interaction between the Tat protein and its target RNA TAR element can also be probed using RNA-ChIP. Still, not all mutations affected binding sites for transcription factors so this is not the only mechanism by which the A3 proteins modulate promoter activity.

Given the vast number of mutations discovered in the LTRs of our *in vitro* clones, we wanted to know if these same mutations were also found in a more representative environment, such as within HIV-1-infected patient sequences. In support of our experimental findings, we were able to identify many of the same mutations and targeted transcription factor binding sites in the LTRs of the patient sequences as in our *in vitro* data. This tells us that our assays are uncovering mutations that appear physiologically in nature and are relevant to study.

One of our goals is to discover precise mutations or a minimum set of mutations that would generate a latent phenotype. I generated SDM clones of potentially interesting mutations in order to test their individual effects. To help me choose which mutations to focus on, I looked at the mean fluorescence intensity of each individual clone at all the mutated positions. Positions that were mutated and consistently displayed weak fluorescence intensities

in all the clones with that mutated position were appealing to me because they were the ones most hopeful to contribute to a latency-prone phenotype. Approaching this goal from the other end, I also chose one of the clones in my mutated library that had a few number of mutations but also displayed a low eGFP fluorescence. To this clone, I individually reverted a number of its G-to-A mutations via SDM back to a guanine residue in effort to decompose this clone until I reached a minimum set of mutations leading to a latent phenotype. The LTRs of all of these clones, as well as the mutated clones from my reporter library, were reconstituted into a non-replicative HIV-1 expression vector. Since only the minimally-mutated single- and double-point mutant clones were able to infect in our pilot infection assay, we decided to delve deeper into why the heavily mutated and revertant clones were non-infectious. We speculated that perhaps certain regions of the LTR, such as the R region, must be intact in order for proper reverse transcription and infection of the virus to take place. Since it is already known that the U3 region contains the majority of the transcription factor binding sites and their ineffective binding may be one of the mechanisms for reduced transcription in the mutated clones, we also wanted to investigate what effects mutations in the R and U5 regions had on the activity of the promoter. I returned to my eGFP reporter system for this and created partially-mutated LTR clones where I took the mutated U3 regions from my reporter library and fused them with wildtype R and U5 regions. I then compared the eGFP expression and fluorescence differences between these partially-mutated clones with their fully-mutated counterparts and found that mutations in the R and U5 regions did have an effect on promoter activity as their absence led to some of the clones expressing better or worse in comparison to the fully-mutated clones. The few mutations in the R and U5 regions may not be individually compensatory or detrimental but rather their overall effect together contributes to this phenotype. Moreover, because the R and U5 regions bear few transcription factor binding sites, this would suggest

once again that other mechanisms of promoter modulation by the A3 proteins exist in addition to affecting transcription factor binding. The mutations can potentially be affecting RNA binding and secondary structure. In a study by Cupelli *et al.* (188), they found that mutations in the secondary structure of the R region affected LTR-driven gene expression in Moloney murine leukemia virus (a gammaretrovirus). Furthermore, we know that the TAR element is located in the R region and as we discovered, it is mutated frequently by the A3s in both our *in vitro* clones as well as in archival HIV-1-infected patient sequences. The TAR element is a stem-loop RNA structure in the viral genome that would be affected by mutations in the R region. The data from these experiments therefore demonstrate once again that many avenues exist for A3-induced promoter modulation. It is clear that further investigation into the exact mechanisms by which the A3 mutations affect promoter activity is warranted.

Despite their impairment in infectivity even with Tat-induction, we wanted to test whether the heavily mutated and revertant reconstituted clones could be activated using latency reversing agents. We chose to test their reactivation potential using PMA and ionomycin first as they are robust activators of T cells. We tested the reactivation potential of both these clone sets as well as the single- and double-point mutant clones with this cocktail and were enthused to find that some of the heavily mutated clones were in fact able to be reactivated with this LRA. Some clones were already able to infect without induction and following stimulation, induced to levels in the range of the wildtype control or within a similarly highly infectious range. Other clones could not be induced to physiologically-relevant detectable levels and were thus poorly infectious or inactivated. Both these types of clones are not of interest to us as they do not represent a latency-prone virus. We initially set our parameters for a latency-prone virus among these clones to be any clone that initially had a basal infection of 0.25% or lower but was able to be induced 10-fold or higher with PMA and ionomycin stimulation. We

were able to uncover 10 such clones with compromised infectivity that matched these criteria. These clones represent viruses that initially, without stimulation, are so weakly-transcribing or inactivated that they go undetected by the host immune system and antiretroviral defenses. Upon stimulation, these latency-prone viruses are able to be reactivated and can contribute to the spread of infection as well as repopulation of the latent reservoir. Although the clones with more mutations generally were more compromised than clones with a few mutations, this is not a finite trend as there were clones where this trend did not fit. It is thus perhaps the location of these mutations that is more important than the number of mutations that an LPV possesses that determines their latent state. We then take a look at our single- and double-point mutant clones. As these clones only differ from the wildtype control by one or two mutations in the LTR of the virus, the majority of these clones were highly infectious. In fact, only 8 clones displayed infectivities noticeably lower than the wildtype control, yet were still too infectious on their own to be considered an LPV. Although these results demonstrate that not one mutation can be responsible for generating a latent phenotype, a combination of these 8 mutations could have the potential to lead to the generation of an LPV. It is also highly favourable that these 8 slightly impaired mutants were also the ones most highly-induced after stimulation, which make them even more interesting as candidate latency-inducing mutations, having the greatest dynamic range after induction. If we look at where these 8 mutations are located, they are within binding sites for the transcription factors NFAT, Sp1, NF-κB, as well as within the *Nef* gene and the TAR element, among others. Since we already know these particular sites are essential for transcription, it may explain why mutations at these spots lead to the decrease in infectivity and MFI observed in these mutants. Additionally, the six double-point mutants were created as combinations of four positions that were mutated most frequently in the heavily mutated clones (positions 350, 393, 454, and 485). Individually, they

do not appear to have a drastic effect on promoter activity, although mutations at positions 393 and 454 do appear to result in lower infection than mutations at positions 350 and 485, which appear to result in higher infection in comparison. Accordingly, the double-point mutant containing mutations at both 393 and 454 was the most compromised of the double-point mutants, with the others behaving as a combined effect of the two mutations they possess. To fully elucidate the repertoire of latency-inducing mutations, SDMs of all A3 deamination target sites could have been generated. However, this may be a costly and timely endeavour and the positions we chose to mutate were selected based on their consistent lowered fluorescence and thus were deemed the most influential positions to mutate. Although a virus containing a combination of these 8 mutations may not exist naturally in the circulating population (or at least were not uncovered in our clones), we still have the added benefit of identifying these mutations in synthetically generating such a virus with these mutations to use as a tool or model to study latency and novel latency-reversing agents. Despite our efforts to work our way backwards from an already weak clone (D9) towards a minimum set of mutations that would generate a latent phenotype, none of the nine positions we chose to mutate were able to rescue the inactivated phenotype of clone D9. This once again demonstrates that the activity of the promoter and entry into a latent state is not dependent on the action of one mutation but rather a combination of mutations. Furthermore, the effects of the eight other mutations in the LTR of clone D9 could have also had an effect but were not explored.

Altogether, the data does support our hypothesis that low levels of A3-induced mutations in the LTR promoter of HIV-1 generate weakly-transcribing and latent viruses in the infected cell. We were able to demonstrate that the HIV-1 LTR is mutated by A3G and A3F and that these mutations affect promoter activity and transactivation by Tat as well as induction by

latency-reversing agents such as PMA and ionomycin. The mutations we have uncovered are located at important transcription factor binding sites as well as within the *Nef* gene which is important early on in HIV-1 infection for T cell activation and the establishment of persistent infection, and in the TAR element which is crucial for Tat binding and efficient transcription. These mutations were also similarly found in HIV-1-infected patient sequences which validates their importance and existence in nature. We were able to shed light into the mechanism of A3 deamination in the HIV-1 promoter and discovered that the 5' LTR is not mutated during the first round of infection but rather the 3' LTR and that two rounds of infection must take place in order to generate the latent viruses by A3. We also further examined the regions of the LTR and found that not only was the TFBS-heavy U3 region important but that mutations in the R and U5 regions also had an effect on promoter activity as well, which warrants further investigation into other methods that these mutations are causing latency by, such as through RNA folding, etc. Finally, we have identified at least 10 latency-prone viruses and at least 8 SDM mutants containing mutations of potential interest. This not only demonstrates for the first time that mutations caused by the APOBEC3 proteins can potentially generate a subset of latent viruses but it also tells us that previous known mechanisms for the establishment and maintenance of latency are not complete. This may explain why current latency-reversing agents do not work on all latent viruses as they mainly target epigenetic mechanisms of latency. This necessitates investigation into the development of novel latency-reversing agents that act in alternate ways as well as the potential development of anti-APOBEC3 agents that may prevent the generation of this subset of latent viruses. However, the A3 proteins make up a huge part of our intrinsic immunity against retroviruses and there needs to be a way to balance their adverse effect with latency and their beneficial

effect in our protection. Nevertheless, the retroviral restriction afforded by the A3 proteins may be compensated by the action of the other restriction factors found within our cells.

We would also like to eventually move our studies into humanized mice models to better reproduce a physiological environment and be able to further study latency *in vivo* as well as test latency-reversing agents with a more intact immune system. Before we get to this stage however, we should first perform experiments to test whether the latency-prone viruses we discovered can propagate *in vivo*. We would clone these LPVs into replicative pNL4.3-GFP_IRES_Nef WT (X4-tropic) and pNL4.3-ADA-GFP_IRES_Nef WT (R5-tropic) reporter viruses and observe how these viruses propagate through multiple rounds of infection in a variety of T cell lines and primary cells. In addition, we should also study the effects of other LRAs on the latency-prone viruses as the combination of PMA and ionomycin is a robust T cell activator and is suitable as a reference but does not represent the performance of other LRAs that may be used in humans. Once again, our LPVs would make an excellent tool for testing existing as well as novel therapeutics.

As mentioned earlier, although I only studied APOBEC3-induced mutations in the LTR of HIV-1, we must not forget that the virus would accumulate mutations throughout the rest of its genome as well. One way that the virus may overcome the sublethal mutations caused by A3 is through viral polymerase error. The poor fidelity of the HIV-1 reverse transcriptase has the capacity to reverse the sublethal mutations generated by APOBEC3. A virus with previous defects could become restored to a more infectious and replicatively fit form through RT error. These viruses could then go on to propagate the infection in the infected host. This would demonstrate that these LPVs are not only reversibly-latent, but also have the potential to lose their latent phenotype altogether over time. This presents another facet to be studied in understanding the mechanisms involved in latency generated by the APOBEC3 proteins.

6.0 CONCLUSION

This thesis demonstrates the ability of A3 proteins to modulate the activity of the HIV-1 LTR promoter through sublethal mutagenesis. We show here for the first time that a subset of latently-infected cells can be generated through A3-induced mutations at regulatory regions of HIV-1. This data supports a novel latency-causing mechanism through molecular modification. This is the first evidence of such a phenomenon as prior to this, the main mechanism for the establishment of latency revolved around epigenetic modifications. This may explain why current latency-reversing agents, which mainly target reactivation and expression of the latent viruses do not work in all patients. The latency-prone viruses we uncovered not only provide a new tool in which to test latency-reversing agents, but also necessitate the development of novel LRAs that target areas other than cell cycle state. The dual role of the A3 proteins in retroviral restriction through hypermutation, and promotion of viral latency through sublethal mutagenesis is clearly illustrated in this work. This brings forth the demand for the development of catalytic inhibitors against the A3 proteins to prevent the latter from occurring. My project demonstrates for the first time a role for A3 deamination in HIV-1 latency and advances the knowledge towards the goal of a functional cure for HIV-1.

7.0 REFERENCES

1. Coffin JM, Hughes SH, Varmus HE. 1997. Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
2. Gallo RC. 1988. HIV--the cause of AIDS: an overview on its biology, mechanisms of disease induction, and our attempts to control it. J Acquir Immune Defic Syndr 1:521–535.
3. Yoshida M, Miyoshi I, Hinuma Y. 1982. Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. Proc Natl Acad Sci USA 79:2031–2035.
4. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vézinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L. 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science (80-) 220:868–871.
5. Eberle J, Gürtler L. 2012. HIV types, groups, subtypes and recombinant forms: errors in replication, selection pressure and quasispecies. Intervirology 55:79–83.
6. Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med 1:a006841.
7. UNAIDS. 2018. Global Statistics. HIV.gov.
8. Vogt VM. 1997. Retroviral Virions and Genomes, p. . *In* Coffin, JM, Hughes, SH, Varmus, HE (eds.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
9. White TA, Bartesaghi A, Borgnia MJ, Meyerson JR, de la Cruz MJ V, Bess JW, Nandwani R, Hoxie JA, Lifson JD, Milne JLS, Subramaniam S. 2010. Molecular architectures of trimeric SIV and HIV-1 envelope glycoproteins on intact viruses: strain-dependent variation in quaternary structure. PLoS Pathog 6:e1001249.
10. Opijnen T van, Jeeninga RE, Boerlijst MC, Pollakis GP, Zetterberg V, Salminen M, Berkhout B. 2004. Human Immunodeficiency Virus Type 1 Subtypes Have a Distinct Long Terminal Repeat That Determines the Replication Rate in a Host-Cell-Specific Manner. J Virol 78:3675–3683.
11. Gelderblom HR, Ozel M, Pauli G. 1989. Morphogenesis and morphology of HIV. Structure-function relations. Arch Virol 106:1–13.
12. Lu K, Heng X, Summers MF. 2011. Structural determinants and mechanism of HIV-1 genome packaging. J Mol Biol 410:609–633.
13. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW, Swanstrom R, Burch CL, Weeks KM. 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature 460:711–716.
14. Suzuki Y, Suzuki Y. 2011. Gene Regulatable Lentiviral Vector System, p. 285–308. *In* Xu, K (ed.), Viral Gene Therapy. InTech.
15. Steckbeck JD, Kuhlmann A-S, Montelaro RC. 2013. C-terminal tail of human immunodeficiency virus gp41: functionally rich and structurally enigmatic. J Gen Virol 94:1–19.
16. Freed EO, Martin MA. 1995. The role of human immunodeficiency virus type 1 envelope glycoproteins in virus infection. J Biol Chem 270:23883–23886.
17. Kim Y, Anderson JL, Lewin SR. 2018. Getting the “Kill” into “Shock and Kill”: Strategies to Eliminate Latent HIV. Cell Host Microbe 23:14–26.
18. Wang X, Ye L, Hou W, Zhou Y, Wang Y-J, Metzger DS, Ho W-Z. 2009. Cellular

microRNA expression correlates with susceptibility of monocytes/macrophages to HIV-1 infection. Blood 113:671–674.
19. Coleman CM, Wu L. 2009. HIV interactions with monocytes and dendritic cells: viral latency and reservoirs. Retrovirology 6:51.
20. Huang CB, Alimova Y V, Ebersole JL. 2011. HIV-1 reactivation in HIV-latently infected dendritic cells by oral microorganisms and LPS. Cell Immunol 268:105–111.
21. Wilen CB, Tilton JC, Doms RW. 2012. HIV: cell binding and entry. Cold Spring Harb Perspect Med 2.
22. Espy N, Pacheco B, Sodroski J. 2017. Adaptation of HIV-1 to cells with low expression of the CCR5 coreceptor. Virology 508:90–107.
23. Jiang X, Feyertag F, Robertson DL. 2017. Protein structural disorder of the envelope V3 loop contributes to the switch in human immunodeficiency virus type 1 cell tropism. PLoS One 12:e0185790.
24. Korber BT, Farber RM, Wolpert DH, Lapedes AS. 1993. Covariation of mutations in the V3 loop of human immunodeficiency virus type 1 envelope protein: an information theoretic analysis. Proc Natl Acad Sci USA 90:7176–7180.
25. Berger EA, Murphy PM, Farber JM. 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. Annu Rev Immunol 17:657–700.
26. Doms RW, Peiper SC. 1997. Unwelcomed guests with master keys: how HIV uses chemokine receptors for cellular entry. Virology 235:179–190.
27. Merk A, Subramaniam S. 2013. HIV-1 envelope glycoprotein structure. Curr Opin Struct Biol 23:268–276.
28. Checkley MA, Luttge BG, Freed EO. 2011. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. J Mol Biol 410:582–608.
29. Ward AB, Wilson IA. 2015. Insights into the trimeric HIV-1 envelope glycoprotein structure. Trends Biochem Sci 40:101–107.
30. Roberts JD, Bebenek K, Kunkel TA. 1988. The accuracy of reverse transcriptase from HIV-1. Science (80-) 242:1171–1173.
31. Liu R, Wu J, Shao R, Xue Y. 2014. Mechanism and factors that control HIV-1 transcription and latency activation. J Zhejiang Univ Sci B 15:455–465.
32. Arhel N. 2010. Revisiting HIV-1 uncoating. Retrovirology 7:96.
33. Telesnitsky A, Goff SP. 1997. Reverse Transcriptase and the Generation of Retroviral DNA, p. . *In* Coffin, JM, Hughes, SH, Varmus, HE (eds.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
34. Zheng Y-H, Lovsin N, Peterlin BM. 2005. Newly identified host factors modulate HIV replication. Immunol Lett 97:225–234.
35. Dahabieh MS, Battivelli E, Verdin E. 2015. Understanding HIV Latency: The Road to an HIV Cure. Annu Rev Med 66:407–421.
36. Brown PO. 1997. Integration, p. . *In* Coffin, JM, Hughes, SH, Varmus, HE (eds.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
37. Hu W-S, Hughes SH. 2012. HIV-1 Reverse Transcription. Cold Spring Harb Perspect Med 2:a006882--a006882.
38. Fuda NJ, Ardehali MB, Lis JT. 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature 461:186–192.
39. Fisher AG, Feinberg MB, Josephs SF, Harper ME, Marselle LM, Reyes G, Gonda MA, Aldovini A, Debouk C, Gallo RC. 1986. The trans-activator gene of HTLV-III is

essential for virus replication. Nature 320:367–371.
40. Dayton AI, Sodroski JG, Rosen CA, Goh WC, Haseltine WA. 1986. The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. Cell 44:941–947.
41. Ruelas DS, Greene WC. 2013. An Integrated Overview of HIV-1 Latency. Cell 155:519–529.
42. Daly TJ, Cook KS, Gray GS, Maione TE, Rusche JR. 1989. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. Nature 342:816–819.
43. Berkowitz R, Fisher J, Goff SP. 1996. RNA packaging. Curr Top Microbiol Immunol 214:177–218.
44. Brik A, Wong C-H. 2003. HIV-1 protease: mechanism and drug discovery. Org Biomol Chem 1:5–14.
45. Wensing AMJ, van Maarseveen NM, Nijhuis M. 2010. Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. Antivir Res 85:59–74.
46. Rambaut A, Posada D, Crandall KA, Holmes EC. 2004. The causes and consequences of HIV evolution. Nat Rev Genet 5:52–61.
47. Russell ES, Kwiek JJ, Keys J, Barton K, Mwapasa V, Montefiori DC, Meshnick SR, Swanstrom R. 2011. The genetic bottleneck in vertical transmission of subtype C HIV-1 is not driven by selection of especially neutralization-resistant virus from the maternal viral population. J Virol 85:8253–8262.
48. Tully DC, Ogilvie CB, Batorsky RE, Bean DJ, Power KA, Ghebremichael M, Bedard HE, Gladden AD, Seese AM, Amero MA, Lane K, McGrath G, Bazner SB, Tinsley J, Lennon NJ, Henn MR, Brumme ZL, Norris PJ, Rosenberg ES, Mayer KH, Jessen H, Kosakovsky Pond SL, Walker BD, Altfeld M, Carlson JM, Allen TM. 2016. Differences in the Selection Bottleneck between Modes of Sexual Transmission Influence the Genetic Composition of the HIV-1 Founder Virus. PLoS Pathog 12:e1005619.
49. Li H, Bar KJ, Wang S, Decker JM, Chen Y, Sun C, Salazar-Gonzalez JF, Salazar MG, Learn GH, Morgan CJ, Schumacher JE, Hraber P, Giorgi EE, Bhattacharya T, Korber BT, Perelson AS, Eron JJ, Cohen MS, Hicks CB, Haynes BF, Markowitz M, Keele BF, Hahn BH, Shaw GM. 2010. High Multiplicity Infection by HIV-1 in Men Who Have Sex with Men. PLoS Pathog 6:e1000890.
50. Coakley E, Petropoulos CJ, Whitcomb JM. 2005. Assessing chemokine co-receptor usage in HIV. Curr Opin Infect Dis 18:9–15.
51. Gurunathan S, Habib R El, Baglyos L, Meric C, Plotkin S, Dodet B, Corey L, Tartaglia J. 2009. Use of predictive markers of HIV disease progression in vaccine trials. Vaccine 27:1997–2015.
52. Douek DC, Roederer M, Koup RA. 2009. Emerging concepts in the immunopathogenesis of AIDS. Annu Rev Med 60:471–484.
53. Mattapallil JJ, Douek DC, Hill B, Nishimura Y, Martin M, Roederer M. 2005. Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. Nature 434:1093–1097.
54. Guadalupe M, Reay E, Sankaran S, Prindiville T, Flamm J, McNeil A, Dandekar S. 2003. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. J Virol 77:11708–11717.
55. Brenchley JM, Schacker TW, Ruff LE, Price DA, Taylor JH, Beilman GJ, Nguyen

PL, Khoruts A, Larson M, Haase AT, Douek DC. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med 200:749–759.
56. Mehandru S, Poles MA, Tenner-Racz K, Horowitz A, Hurley A, Hogan C, Boden D, Racz P, Markowitz M. 2004. Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. J Exp Med 200:761–770.
57. Lewin-Smith MR, Klassen MK, Frankel SS, Nelson AM. 1998. Pathology of human immunodeficiency virus infection: infectious conditions. Ann Diagn Pathol 2:181–194.
58. Collier AC, Coombs RW, Schoenfeld DA, Bassett RL, Timpone J, Baruch A, Jones M, Facey K, Whitacre C, McAuliffe VJ, Friedman HM, Merigan TC, Reichman RC, Hooper C, Corey L. 1996. Treatment of human immunodeficiency virus infection with saquinavir, zidovudine, and zalcitabine. AIDS Clinical Trials Group. N Engl J Med 334:1011–1017.
59. D'Aquila RT, Hughes MD, Johnson VA, Fischl MA, Sommadossi JP, Liou SH, Timpone J, Myers M, Basgoz N, Niu M, Hirsch MS. 1996. Nevirapine, zidovudine, and didanosine compared with zidovudine and didanosine in patients with HIV-1 infection. A randomized, double-blind, placebo-controlled trial. National Institute of Allergy and Infectious Diseases AIDS Clinical Trials Group Protoco. Ann Intern Med 124:1019–1030.
60. Staszewski S, Miller V, Rehmet S, Stark T, De Crée J, De Brabander M, Peeters M, Andries K, Moeremans M, De Raeymaeker M, Pearce G, den Broeck R, Verbiest W, Stoffels P. 1996. Virological and immunological analysis of a triple combination pilot study with loviride, lamivudine and zidovudine in HIV-1-infected patients. AIDS 10:F1--7.
61. Pau AK, George JM. 2014. Antiretroviral therapy: current drugs. Infect Dis Clin North Am 28:371–402.
62. Coffin JM. 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science (80-) 267:483–489.
63. Abram ME, Ferris AL, Shao W, Alvord WG, Hughes SH. 2010. Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. J Virol 84:9864–9878.
64. Abram ME, Ferris AL, Das K, Quinoñes O, Shao W, Tuske S, Alvord WG, Arnold E, Hughes SH. 2014. Mutations in HIV-1 reverse transcriptase affect the errors made in a single cycle of viral replication. J Virol 88:7589–7601.
65. Cuevas JM, Geller R, Garijo R, López-Aldeguer J, Sanjuán R. 2015. Extremely High Mutation Rate of HIV-1 In Vivo. PLoS Biol 13:e1002251.
66. Andrews SM, Rowland-Jones S. 2017. Recent advances in understanding HIV evolution. F1000Res 6:597.
67. Mansky LM, Temin HM. 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J Virol 69:5087–5094.
68. O'Neil PK, Sun G, Yu H, Ron Y, Dougherty JP, Preston BD. 2002. Mutational analysis of HIV-1 long terminal repeats to explore the relative contribution of reverse transcriptase and RNA polymerase II to viral mutagenesis. J Biol Chem 277:38053–38061.

69. Autran B, Carcelain G, Li TS, Blanc C, Mathez D, Tubiana R, Katlama C, Debré P, Leibowitch J. 1997. Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease. Science (80-) 277:112–116.
70. Komanduri K V, Viswanathan MN, Wieder ED, Schmidt DK, Bredt BM, Jacobson MA, McCune JM. 1998. Restoration of cytomegalovirus-specific CD4+ T-lymphocyte responses after ganciclovir and highly active antiretroviral therapy in individuals infected with HIV-1. Nat Med 4:953–956.
71. Lederman MM, Connick E, Landay A, Kuritzkes DR, Spritzler J, St Clair M, Kotzin BL, Fox L, Chiozzi MH, Leonard JM, Rousseau F, Wade M, Roe JD, Martinez A, Kessler H. 1998. Immunologic responses associated with 12 weeks of combination antiretroviral therapy consisting of zidovudine, lamivudine, and ritonavir: results of AIDS Clinical Trials Group Protocol 315. J Infect Dis 178:70–79.
72. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, Smith K, Lisziewicz J, Lori F, Flexner C, Quinn TC, Chaisson RE, Rosenberg E, Walker B, Gange S, Gallant J, Siliciano RF. 1999. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. Nat Med 5:512–517.
73. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, Kovacs C, Gange SJ, Siliciano RF. 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nat Med 9:727–728.
74. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, Quinn TC, Chadwick K, Margolick J, Brookmeyer R, Gallant J, Markowitz M, Ho DD, Richman DD, Siliciano RF. 1997. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. Science (80-) 278:1295–1300.
75. Rouzine IM, Weinberger AD, Weinberger LS. 2015. An Evolutionary Role for HIV Latency in Enhancing Viral Transmission. Cell 160:1002–1012.
76. Razooky BS, Pai A, Aull K, Rouzine IM, Weinberger LS. 2015. A Hardwired HIV Latency Program. Cell 160:990–1001.
77. Ho Y-C, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DIS, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-Competent Noninduced Proviruses in the Latent Reservoir Increase Barrier to HIV-1 Cure. Cell 155:540–551.
78. Shan L, Deng K, Gao H, Xing S, Capoferri AA, Durand CM, Rabi SA, Laird GM, Kim M, Hosmane NN, Yang H-C, Zhang H, Margolick JB, Li L, Cai W, Ke R, Flavell RA, Siliciano JD, Siliciano RF. 2017. Transcriptional Reprogramming during Effector-to-Memory Transition Renders CD4+ T Cells Permissive for Latent HIV-1 Infection. Immunity 47:766--775.e3.
79. Coiras M, López-Huertas MR, Pérez-Olmeda M, Alcamí J. 2009. Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. Nat Rev Microbiol 7:798–812.
80. Eisele E, Siliciano RF. 2012. Redefining the Viral Reservoirs that Prevent HIV-1 Eradication. Immunity 37:377–388.
81. Cohn LB, Silva IT, Oliveira TY, Rosales RA, Parrish EH, Learn GH, Hahn BH, Czartoski JL, McElrath MJ, Lehmann C, Klein F, Caskey M, Walker BD, Siliciano JD, Siliciano RF, Jankovic M, Nussenzweig MC. 2015. HIV-1 Integration Landscape during Latent and Active Infection. Cell 160:420–432.
82. Bullen CK, Laird GM, Durand CM, Siliciano JD, Siliciano RF. 2014. New ex vivo approaches distinguish effective and ineffective single agents for reversing HIV-1

latency in vivo. Nat Med 20:425–429.
83. Chavez L, Calvanese V, Verdin E. 2015. HIV Latency Is Established Directly and Early in Both Resting and Activated Primary CD4 T Cells. PLoS Pathog 11:e1004955.
84. Chun TW, Carruth L, Finzi D, Shen X, DiGiuseppe JA, Taylor H, Hermankova M, Chadwick K, Margolick J, Quinn TC, Kuo YH, Brookmeyer R, Zeiger MA, Barditch-Crovo P, Siliciano RF. 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature 387:183–188.
85. Cary DC, Fujinaga K, Peterlin BM. 2016. Molecular mechanisms of HIV latency. J Clin Invest 126:448–454.
86. Tsang J, Chain BM, Miller RF, Webb BLJ, Barclay W, Towers GJ, Katz DR, Noursadeghi M. 2009. HIV-1 infection of macrophages is dependent on evasion of innate immune cellular activation. AIDS 23:2255–2263.
87. Honeycutt JB, Thayer WO, Baker CE, Ribeiro RM, Lada SM, Cao Y, Cleary RA, Hudgens MG, Richman DD, Garcia JV. 2017. HIV persistence in tissue macrophages of humanized myeloid-only mice during antiretroviral therapy. Nat Med 23:638–643.
88. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim E-Y, Archer J, Pond SLK, Chung Y-S, Penugonda S, Chipman J, Fletcher C V, Schacker TW, Malim MH, Rambaut A, Haase AT, McLean AR, Wolinsky SM. 2016. Persistent HIV-1 replication maintains the tissue reservoir during therapy. Nature 530:51–56.
89. Joseph SB, Arrildt KT, Sturdevant CB, Swanstrom R. 2015. HIV-1 target cells in the CNS. J Neurovirol 21:276–289.
90. Kandathil AJ, Sugawara S, Balagopal A. 2016. Are T cells the only HIV-1 reservoir? Retrovirology 13:86.
91. Margolis DM, Garcia JV, Hazuda DJ, Haynes BF. 2016. Latency reversal and viral clearance to cure HIV-1. Science (80-) 353:aaf6517.
92. Van Lint C, Bouchat S, Marcello A. 2013. HIV-1 transcription and latency: an update. Retrovirology 10:67.
93. Mbonye U, Karn J. 2017. The Molecular Basis for Human Immunodeficiency Virus Latency. Annu Rev Virol 4:261–285.
94. Spivak AM, Planelles V. 2018. Novel Latency Reversal Agents for HIV-1 Cure. Annu Rev Med 69:421–436.
95. Rasmussen TA, Schmeltz Søgaard O, Brinkmann C, Wightman F, Lewin SR, Melchjorsen J, Dinarello C, Østergaard L, Tolstrup M. 2013. Comparison of HDAC inhibitors in clinical development: effect on HIV production in latently infected cells and T-cell activation. Hum Vaccin Immunother 9:993–1001.
96. Margolis DM. 2011. Histone deacetylase inhibitors and HIV latency. Curr Opin HIV AIDS 6:25–29.
97. McKernan LN, Momjian D, Kulkosky J. 2012. Protein Kinase C: One Pathway towards the Eradication of Latent HIV-1 Reservoirs. Adv Virol 2012:805347.
98. Marsden MD, Zack JA. 2014. Neutralizing the HIV Reservoir. Cell 158:971–972.
99. Spivak AM, Planelles V. 2016. HIV-1 Eradication: Early Trials (and Tribulations). Trends Mol Med 22:10–27.
100. Katlama C, Deeks SG, Autran B, Martinez-Picado J, van Lunzen J, Rouzioux C, Miller M, Vella S, Schmitz JE, Ahlers J, Richman DD, Sekaly RP. 2013. Barriers to a cure for HIV: new ways to target and eradicate HIV-1 reservoirs. Lancet 381:2109–2117.

101. Melkova Z, Shankaran P, Madlenakova M, Bodor J. 2017. Current views on HIV-1 latency, persistence, and cure. Folia Microbiol 62:73–87.
102. Nisole S, Stoye JP, Saïb A. 2005. TRIM family proteins: retroviral restriction and antiviral defence. Nat Rev Microbiol 3:799–808.
103. Towers GJ. 2007. The control of viral infection by tripartite motif proteins and cyclophilin A. Retrovirology 4:40.
104. Stremlau M, Perron M, Lee M, Li Y, Song B, Javanbakht H, Diaz-Griffero F, Anderson DJ, Sundquist WI, Sodroski J. 2006. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. Proc Natl Acad Sci USA 103:5514–5519.
105. Azevedo-Pereira JM, Canhão P, Calado M, Santos- Costa Q, Barroca P. 2015. Inhibition of HIV Replication by Host Cellular Factors. Trends Basic Ther Options HIV Infect - Towar a Funct Cure.
106. Strebel K. 2013. HIV accessory proteins versus host restriction factors. Curr Opin Virol 3:692–699.
107. Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J. 2004. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature 427:848–853.
108. Kirmaier A, Wu F, Newman RM, Hall LR, Morgan JS, O'Connor S, Marx PA, Meythaler M, Goldstein S, Buckler-White A, Kaur A, Hirsch VM, Johnson WE. 2010. TRIM5 suppresses cross-species transmission of a primate immunodeficiency virus and selects for emergence of resistant variants in the new species. PLoS Biol 8.
109. Li N, Zhang W, Cao X. 2000. Identification of human homologue of mouse IFN-gamma induced protein from human dendritic cells. Immunol Lett 74:221–224.
110. Skasko M, Weiss KK, Reynolds HM, Jamburuthugoda V, Lee K, Kim B. 2005. Mechanistic differences in RNA-dependent DNA polymerization and fidelity between murine leukemia virus and HIV-1 reverse transcriptases. J Biol Chem 280:12190–12200.
111. Diamond TL, Roshal M, Jamburuthugoda VK, Reynolds HM, Merriam AR, Lee KY, Balakrishnan M, Bambara RA, Planelles V, Dewhurst S, Kim B. 2004. Macrophage tropism of HIV-1 depends on efficient cellular dNTP utilization by reverse transcriptase. J Biol Chem 279:51545–51553.
112. Baldauf H-M, Pan X, Erikson E, Schmidt S, Daddacha W, Burggraf M, Schenkova K, Ambiel I, Wabnitz G, Gramberg T, Panitz S, Flory E, Landau NR, Sertel S, Rutsch F, Lasitschka F, Kim B, König R, Fackler OT, Keppler OT. 2012. SAMHD1 restricts HIV-1 infection in resting CD4+ T cells. Nat Med 18:1682–1687.
113. Descours B, Cribier A, Chable-Bessia C, Ayinde D, Rice G, Crow Y, Yatim A, Schwartz O, Laguette N, Benkirane M. 2012. SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4+ T-cells. Retrovirology 9:87.
114. Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, Bloch N, Maudet C, Bertrand M, Gramberg T, Pancino G, Priet S, Canard B, Laguette N, Benkirane M, Transy C, Landau NR, Kim B, Margottin-Goguet F. 2012. SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. Nat Immunol 13:223–228.
115. Kupzig S, Korolchuk V, Rollason R, Sugden A, Wilde A, Banting G. 2003. Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology. Traffic 4:694–709.

116. Neil SJD, Sandrin V, Sundquist WI, Bieniasz PD. 2007. An interferon-alpha-induced tethering mechanism inhibits HIV-1 and Ebola virus particle release but is counteracted by the HIV-1 Vpu protein. Cell Host Microbe 2:193–203.
117. Neil SJD, Zang T, Bieniasz PD. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature 451:425–430.
118. Smith HC, Bennett RP, Kizilyer A, McDougall WM, Prohaska KM. 2012. Functions and regulation of the APOBEC family of proteins. Semin Cell Dev Biol 23:258–268.
119. Refsland EW, Harris RS. 2013. The APOBEC3 Family of Retroelement Restriction Factors, p. 1–27. *In* Cullen, BR (ed.), Intrinsic Immunity. Springer Berlin Heidelberg, Berlin, Heidelberg.
120. Vieira VC, Soares MA. 2013. The role of cytidine deaminases on innate immune responses against human viral infections. Biomed Res Int 2013:683095.
121. Hultquist JF, Lengyel JA, Refsland EW, LaRue RS, Lackey L, Brown WL, Harris RS. 2011. Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. J Virol 85:11220–11234.
122. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, Lai J, Strain MC, Lada SM, Hoh R, Ho Y-C, Richman DD, Deeks SG, Siliciano JD, Siliciano RF. 2016. Defective proviruses rapidly accumulate during acute HIV-1 infection. Nat Med 22:1043–1049.
123. Willems L, Gillet NA. 2015. APOBEC3 Interference during Replication of Viral Genomes. Viruses 7:2999–3018.
124. Harris RS, Dudley JP. 2015. APOBECs and virus restriction. Virology 479–480:131–145.
125. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. 2010. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. Nucleic Acids Res 38:4274–4284.
126. Chelico L, Pham P, Calabrese P, Goodman MF. 2006. APOBEC3G DNA deaminase acts processively 3' -->5' on single-stranded DNA. Nat Struct Mol Biol 13:392–399.
127. Langlois M-A, Beale RCL, Conticello SG, Neuberger MS. 2005. Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. Nucleic Acids Res 33:1913–1923.
128. Rausch JW, Chelico L, Goodman MF, Le Grice SFJ. 2009. Dissecting APOBEC3G substrate specificity by nucleoside analog interference. J Biol Chem 284:7047–7058.
129. Sheehy AM, Gaddis NC, Choi JD, Malim MH. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 418:646–650.
130. Conticello SG, Thomas CJF, Petersen-Mahrt SK, Neuberger MS. 2005. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Mol Biol Evol 22:367–377.
131. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, Scott J, Navaratnam N. 2002. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics 79:285–296.
132. Haché G, Liddament MT, Harris RS. 2005. The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. J Biol Chem 280:10920–10924.

133. Navarro F, Bollman B, Chen H, König R, Yu Q, Chiles K, Landau NR. 2005. Complementary function of the two catalytic domains of APOBEC3G. Virology 333:374–386.
134. Siu KK, Sultana A, Azimi FC, Lee JE. 2013. Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. Nat Commun 4:2593.
135. Bohn M-F, Shandilya SMD, Albin JS, Kouno T, Anderson BD, McDougle RM, Carpenter MA, Rathore A, Evans L, Davis AN, Zhang J, Lu Y, Somasundaran M, Matsuo H, Harris RS, Schiffer CA. 2013. Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. Structure 21:1042–1050.
136. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH. 2003. DNA deamination mediates innate immunity to retroviral infection. Cell 113:803–809.
137. Liddament MT, Brown WL, Schumacher AJ, Harris RS. 2004. APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. Curr Biol 14:1385–1391.
138. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 424:99–103.
139. Yu Q, König R, Pillai S, Chiles K, Kearney M, Palmer S, Richman D, Coffin JM, Landau NR. 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nat Struct Mol Biol 11:435–442.
140. Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L. 2003. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature 424:94–98.
141. Mariani R, Chen D, Schröfelbauer B, Navarro F, König R, Bollman B, Münk C, Nymark-McMahon H, Landau NR. 2003. Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. Cell 114:21–31.
142. Feng Y, Baig TT, Love RP, Chelico L. 2014. Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. Front Microbiol 5:450.
143. Desimmie BA, Delviks-Frankenberrry KA, Burdick RC, Qi D, Izumi T, Pathak VK. 2014. Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. J Mol Biol 426:1220–1245.
144. Dubé M, Paquay C, Roy BB, Bego MG, Mercier J, Cohen EA. 2011. HIV-1 Vpu antagonizes BST-2 by interfering mainly with the trafficking of newly synthesized BST-2 to the cell surface. Traffic 12:1714–1729.
145. Ayinde D, Maudet C, Transy C, Margottin-Goguet F. 2010. Limelight on two HIV/SIV accessory proteins in macrophage infection: is Vpx overshadowing Vpr? Retrovirology 7:35.
146. Holmes RK, Koning FA, Bishop KN, Malim MH. 2007. APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. J Biol Chem 282:2587–2595.
147. Newman ENC, Holmes RK, Craig HM, Klein KC, Lingappa JR, Malim MH, Sheehy AM. 2005. Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. Curr Biol 15:166–170.
148. Bishop KN, Holmes RK, Malim MH. 2006. Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. J Virol 80:8450–8458.

149. Browne EP, Allers C, Landau NR. 2009. Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent. Virology 387:313–321.
150. Holmes RK, Malim MH, Bishop KN. 2007. APOBEC-mediated viral restriction: not simply editing? Trends Biochem Sci 32:118–128.
151. Kobayashi T, Koizumi Y, Takeuchi JS, Misawa N, Kimura Y, Morita S, Aihara K, Koyanagi Y, Iwami S, Sato K. 2014. Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. J Virol 88:5881–5887.
152. Guo F, Cen S, Niu M, Yang Y, Gorelick RJ, Kleiman L. 2007. The interaction of APOBEC3G with human immunodeficiency virus type 1 nucleocapsid inhibits tRNA3Lys annealing to viral RNA. J Virol 81:11322–11331.
153. Adolph MB, Webb J, Chelico L. 2013. Retroviral restriction factor APOBEC3G delays the initiation of DNA synthesis by HIV-1 reverse transcriptase. PLoS One 8:e64196.
154. Bishop KN, Verma M, Kim E-Y, Wolinsky SM, Malim MH. 2008. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS Pathog 4:e1000231.
155. Iwatani Y, Chan DSB, Wang F, Maynard KS, Sugiura W, Gronenborn AM, Rouzina I, Williams MC, Musier-Forsyth K, Levin JG. 2007. Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. Nucleic Acids Res 35:7096–7108.
156. Wang X, Ao Z, Chen L, Kobinger G, Peng J, Yao X. 2012. The cellular antiviral protein APOBEC3G interacts with HIV-1 reverse transcriptase and inhibits its function during viral replication. J Virol 86:3777–3786.
157. Chaurasiya KR, McCauley MJ, Wang W, Qualley DF, Wu T, Kitamura S, Geertsema H, Chan DSB, Hertz A, Iwatani Y, Levin JG, Musier-Forsyth K, Rouzina I, Williams MC. 2014. Oligomerization transforms human APOBEC3G from an efficient enzyme to a slowly dissociating nucleic acid-binding protein. Nat Chem 6:28–33.
158. Luo K, Wang T, Liu B, Tian C, Xiao Z, Kappes J, Yu X-F. 2007. Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency virus type 1 integrase and inhibit proviral DNA formation. J Virol 81:7238–7248.
159. Haché G, Shindo K, Albin JS, Harris RS. 2008. Evolution of HIV-1 isolates that use a novel Vif-independent mechanism to resist restriction by human APOBEC3G. Curr Biol 18:819–824.
160. Russell RA, Moore MD, Hu W-S, Pathak VK. 2009. APOBEC3G induces a hypermutation gradient: purifying selection at multiple steps during HIV-1 replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA. Retrovirology 6:16.
161. Xu H, Chertova E, Chen J, Ott DE, Roser JD, Hu W-S, Pathak VK. 2007. Stoichiometry of the antiviral protein APOBEC3G in HIV-1 virions. Virology 360:247–256.
162. Venkatesan S, Rosenthal R, Kanu N, McGranahan N, Bartek J, Quezada SA, Hare J, Harris RS, Swanton C. 2018. Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. Ann Oncol 29:563–572.
163. Monajemi M, Woodworth CF, Benkaroun J, Grant M, Larijani M. 2012. Emerging complexities of APOBEC3G action on immunity and viral fitness during HIV infection and treatment. Retrovirology 9:35.
164. Sadler HA, Stenglein MD, Harris RS, Mansky LM. 2010. APOBEC3G contributes to

HIV-1 variation through sublethal mutagenesis. J Virol 84:7396–7404.
165. Kim E-Y, Lorenzo-Redondo R, Little SJ, Chung Y-S, Phalora PK, Maljkovic Berry I, Archer J, Penugonda S, Fischer W, Richman DD, Bhattacharya T, Malim MH, Wolinsky SM. 2014. Human APOBEC3 induced mutation of human immunodeficiency virus type-1 contributes to adaptation and evolution in natural infection. PLoS Pathog 10:e1004281.
166. Jern P, Russell RA, Pathak VK, Coffin JM. 2009. Likely role of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance. PLoS Pathog 5:e1000367.
167. Armitage AE, Deforche K, Welch JJ, Van Laethem K, Camacho R, Rambaut A, Iversen AKN. 2014. Possible footprints of APOBEC3F and/or other APOBEC3 deaminases, but not APOBEC3G, on HIV-1 from patients with acute/early and chronic infections. J Virol 88:12882–12894.
168. Sato K, Takeuchi JS, Misawa N, Izumi T, Kobayashi T, Kimura Y, Iwami S, Takaori-Kondo A, Hu W-S, Aihara K, Ito M, An DS, Pathak VK, Koyanagi Y. 2014. APOBEC3D and APOBEC3F potently promote HIV-1 diversification and evolution in humanized mouse model. PLoS Pathog 10:e1004453.
169. Grant M, Larijani M. 2017. Evasion of adaptive immunity by HIV through the action of host APOBEC3G/F enzymes. AIDS Res Ther 14:44.
170. Wood N, Bhattacharya T, Keele BF, Giorgi E, Liu M, Gaschen B, Daniels M, Ferrari G, Haynes BF, McMichael A, Shaw GM, Hahn BH, Korber B, Seoighe C. 2009. HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. PLoS Pathog 5:e1000414.
171. Ikeda T, Symeonides M, Albin JS, Li M, Thali M, Harris RS. 2018. HIV-1 adaptation studies reveal a novel Env-mediated homeostasis mechanism for evading lethal hypermutation by APOBEC3G. PLoS Pathog 14:e1007010.
172. Bélanger K, Langlois M-A. 2015. Comparative analysis of the gene-inactivating potential of retroviral restriction factors APOBEC3F and APOBEC3G. J Gen Virol 96:2878–2887.
173. Feng Y, Love RP, Chelico L. 2013. HIV-1 viral infectivity factor (Vif) alters processive single-stranded DNA scanning of the retroviral restriction factor APOBEC3G. J Biol Chem 288:6083–6094.
174. Bélanger K, Savoie M, Aydin H, Renner TM, Montazeri Z, Langlois M-A. 2014. Deamination intensity profiling of human APOBEC3 protein activity along the near full-length genomes of HIV-1 and MoMLV by HyperHRM analysis. Virology 448:168–175.
175. Bélanger K, Savoie M, Rosales Gerpe MC, Couture J-F, Langlois M-A. 2013. Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. Nucleic Acids Res 41:7438–7452.
176. Cohen GB, Gandhi RT, Davis DM, Mandelboim O, Chen BK, Strominger JL, Baltimore D. 1999. The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. Immunity 10:661–671.
177. Collins KL, Chen BK, Kalams SA, Walker BD, Baltimore D. 1998. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature 391:397–401.
178. Westervelt P, Henkel T, Trowbridge DB, Orenstein J, Heuser J, Gendelman HE,

Ratner L. 1992. Dual Regulation of Silent and Productive Infection in Monocytes by Distinct Human Immunodeficiency Virus Type 1 Determinants. J VIROL 66:7.
179. Richard J, Sindhu S, Pham TNQ, Belzile J-P, Cohen EA. 2010. HIV-1 Vpr up-regulates expression of ligands for the activating NKG2D receptor and promotes NK cell-mediated killing. Blood 115:1354–1363.
180. Knoepfel SA, Salisch NC, Huelsmann PM, Rauch P, Walter H, Metzner KJ. 2008. Comparison of G-to-A Mutation Frequencies Induced by APOBEC3 Proteins in H9 Cells and Peripheral Blood Mononuclear Cells in the Context of Impaired Processivities of Drug-Resistant Human Immunodeficiency Virus Type 1 Reverse Transcriptase Variants. J Virol 82:6536–6545.
181. Cullen BR. 1991. Regulation of human immunodeficiency virus replication. Annu Rev Microbiol 45:219–250.
182. Gaynor R. 1992. Cellular transcription factors involved in the regulation of HIV-1 gene expression. AIDS 6:347–363.
183. Jones KA, Peterlin BM. 1994. Control of RNA initiation and elongation at the HIV-1 promoter. Annu Rev Biochem 63:717–743.
184. Kingsman SM, Kingsman AJ. 1996. The regulation of human immunodeficiency virus type-1 gene expression. Eur J Biochem 240:491–507.
185. Pereira LA, Bentley K, Peeters A, Churchill MJ, Deacon NJ. 2000. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. Nucleic Acids Res 28:663–668.
186. Ai W, Li H, Song N, Li L, Chen H. 2013. Optimal Method to Stimulate Cytokine Production and Its Use in Immunotoxicity Assessment. Int J Env Res Public Heal 10:3834–3842.
187. Pillai S, Dasgupta P, Chellappan SP. 2009. Chromatin Immunoprecipitation Assays: Analyzing Transcription Factor Binding and Histone Modifications In Vivo, p. 323–339. *In* Chellappan, SP (ed.), Chromatin Protocols: Second Edition. Humana Press, Totowa, NJ.
188. Cupelli L, Okenquist SA, Trubetskoy A, Lenz J. 1998. The Secondary Structure of the R Region of a Murine Leukemia Virus Is Important for Stimulation of Long Terminal Repeat-Driven Gene Expression. J Virol 72:7807–7814.

8.0 APPENDIX

8.1 List of Primers

Table A1. List of primers.

PRIMER NAME	SEQUENCE (5' to 3')
	Primers for Cloning
HcRedC1-Tat SacII FWD	GACTGACTCCGCGG*CT*ATGGAACCAGTCGACCCTAGACTG
HcRedC1-Tat BamHI REV	GACTGACTGGATCCTTAGTGATGGTGATGGTGATGAGATC
HIV LTR FWD deg 2	GCACTAAATCGGAACCCTAGATTAATTRGAARRGCTAATTTRGTCCC
HIV LTR REV 1	GATGCACACAATAGAGGACT
HIV LTR REV 3 deg	GCACTATATCGCAACCCTGAGCTAGCCTGCTAGAGATTTTCYAC
HIV 3' LTR FWD 1	GAATGAGACGAGCTGAGCCAGCAG
HIV 3' LTR REV deg	GCACTATATCGCAACCCTGAGCTAGCTGCTAGAGATTTTCYACACTG
HIV 3' LTR FWD 2 deg	GCACTAAATCGGAACCCTAGATTAATTRGAARRGCTAATTCACTCCC
pNL4-3 XhoI 3' LTR FWD	GACTGACTCTCGAGACCTAGAAAAACATGGAG
pNL4-3 NaeI-SacII 3' LTR REV	GACTGACTGCCGGCGAATTCCCGCGGGTCCCCCCTTTTCTTTTA
peGFP-C3 SacII 3' LTR FWD deg	GACTGACTCCGCGGTRRAARRGCTAATTCACTC
peGFP-C3 NaeI 3' LTR REV deg	GACTGACTGCCGGCTGCTAGAGATTTTYYACAC
HIV 3' LTR R-U5 wt FWD	GACTGACTATTAATGAATTCGCGGCCGCGGTCTCTCTGGTTAGACCAGAT
HIV 3' LTR R-U5 wt REV	GACTGACTGCTAGCTGCTAGAGATTTTCCACACTGAC
peGFP-C3 U3 mut FWD	CCGCCATGCATTAGTTATTAAT
peGFP-C3 U3 mut REV	GACTGACTGCGGCCGCCAGTACAGGCAAAAAGCAGCT
eGFP 2nd Round FWD	ACAACAGCCACAACGTCTATATCAT
293 Stables REV New	CCTCTACAAATGTGGTATGGCTGATTATG
	Primers for Sequencing
Seq Tat in HcRedC1 FWD	CGAGAAGGCCAACTCCGG

Seq Tat in HcRedC1 REV	AGGGGGAGGTGTGGGAGG
Seq peGFP-C3 LTR FWD	TGCTCACATGTTCTTTCCTGCG
Seq peGFP-C3 LTR REV	ACTTGTGGCCGTTTACGTCGCC
Seq HIV int-mut 3' LTR FWD	TGATTGGATGGCCTGCTGTAAG
Seq HIV mut 3' LTR FWD	CCTCAGGTACCTTTAAGACCAATGAC
Seq HIV mut 3' LTR REV	ACTCTGATCCCCGAATGCTTG
Seq HIV 3' LTR SDM FWD	TTCCAGTCACACCTCAGGTAC
Seq HIV 3' LTR SDM REV	AGGTGTGAGCCACTGCGCC
	Primers for Site-Directed Mutagenesis (SDM) *[See Note]
HIV 3' LTR 129 SDM FWD	ATATCCACTGACCTTT**A**GATGGTGCTACAAGCT
HIV 3' LTR 129 SDM REV	AGCTTGTAGCACCATC**T**AAAGGTCAGTGGATAT
HIV 3' LTR 133 SDM FWD	CCACTGACCTTTGGAT**A**GTGCTACAAGCTAGTA
HIV 3' LTR 133 SDM REV	TACTAGCTTGTAGCAC**T**ATCCAAAGGTCAGTGG
HIV 3' LTR 167 SDM FWD	AGTTGAGCCAGATAA**A**GTAGAAGAGGCCAAT
HIV 3' LTR 167 SDM REV	ATTGGCCTCTTCTAC**T**TTATCTGGCTCAACT
HIV 3' LTR A3F SDM 174 FWD	GCCAGATAAGGTAGAA**A**AGGCCAATAAAGGAGA
HIV 3' LTR A3F SDM 174 REV	TCTCCTTTATTGGCCT**T**TTCTACCTTATCTGGC
HIV 3' LTR 186 SDM FWD	AGAAGAGGCCAATAAA**A**GAGAGAACACCAGCTT
HIV 3' LTR 186 SDM REV	AAGCTGGTGTTCTCTC**T**TTTATTGGCCTCTTCT
HIV 3' LTR 189 SDM FWD	AGAGGCCAATAAAGGA**A**AGAACACCAGCTTGTT
HIV 3' LTR 189 SDM REV	AACAAGCTGGTGTTCT**T**TCCTTTATTGGCCTCT
HIV 3' LTR A3F SDM 191 FWD	AGGCCAATAAAGGAGA**A**AACACCAGCTTGTTAC
HIV 3' LTR A3F SDM 191 REV	GTAACAAGCTGGTGTT**T**TCTCCTTTATTGGCCT
HIV 3' LTR 215 SDM FWD	GCTTGTTACACCCTGT**A**AGCCTGCATGGAATGG
HIV 3' LTR 215 SDM REV	CCATTCCATGCAGGCT**T**ACAGGGTGTAACAAGC

HIV 3’ LTR 225 SDM FWD	CCCTGTGAGCCTGCAT**A**GAATGGATGACCCTGA
HIV 3’ LTR 225 SDM REV	TCAGGGTCATCCATTC**T**ATGCAGGCTCACAGGG
HIV 3’ LTR 242 SDM FWD	GAATGGATGACCCTGA**A**AGAGAAGTGTTAGAGT
HIV 3’ LTR 242 SDM REV	ACTCTAACACTTCTCT**T**TCAGGGTCATCCATTC
HIV 3’ LTR 262 SDM FWD	GAAGTGTTAGAGTGGA**A**GTTTGACAGCCGCCTA
HIV 3’ LTR 262 SDM REV	TAGGCGGCTGTCAAAC**T**TCCACTCTAACACTTC
HIV 3’ LTR 312 SDM FWD	CCGAGAGCTGCATCCG**A**AGTACTTCAAGAACTG
HIV 3’ LTR 312 SDM REV	CAGTTCTTGAAGTACT**T**CGGATGCAGCTCTCGG
HIV 3’ LTR 337 SDM FWD	AAGAACTGCTGACATC**A**AGCTTGCTACAAGGGA
HIV 3’ LTR 337 SDM REV	TCCCTTGTAGCAAGCT**T**GATGTCAGCAGTTCTT
HIV 3’ LTR 350 SDM FWD	ATCGAGCTTGCTACAA**A**GGACTTTCCGCTGGGG
HIV 3’ LTR 350 SDM REV	CCCCAGCGGAAAGTCC**T**TTGTAGCAAGCTCGAT
HIV 3’ LTR 351 SDM FWD	TCGAGCTTGCTACAAG**A**GACTTTCCGCTGGGGA
HIV 3’ LTR 351 SDM REV	TCCCCAGCGGAAAGTC**T**CTTGTAGCAAGCTCGA
HIV 3’ LTR 363 SDM FWD-NEW	AAGGGACTTTCCGCT**A**GGGACTTTCCAGGGAGG
HIV 3’ LTR 363 SDM REV-NEW	CCTCCCTGGAAAGTCCC**T**AGCGGAAAGTCCCTT
HIV 3’ LTR 364 SDM FWD	AAGGGACTTTCCGCTG**A**GGACTTTCCAGGGAGG
HIV 3’ LTR 364 SDM REV	CCTCCCTGGAAAGTCC**T**CAGCGGAAAGTCCCTT
HIV 3’ LTR 365 SDM FWD	GGGACTTTCCGCTGG**A**GACTTTCCAGGGAGG
HIV 3’ LTR 365 SDM REV	CCTCCCTGGAAAGTC**T**CCAGCGGAAAGTCCC
HIV 3’ LTR 366 SDM FWD	GGGACTTTCCGCTGGG**A**ACTTTCCAGGGAGGCG
HIV 3’ LTR 366 SDM REV	CGCCTCCCTGGAAAGT**T**CCCAGCGGAAAGTCCC
HIV 3’ LTR 375 SDM FWD	CGCTGGGGACTTTCCA**A**GGAGGCGTGGCCTGGG
HIV 3’ LTR 375 SDM REV	CCCAGGCCACGCCTCC**T**TGGAAAGTCCCCAGCG
HIV 3’ LTR 376 SDM FWD	GCTGGGGACTTTCCAG**A**GAGGCGTGGCCTGGGC

HIV 3’ LTR 376 SDM REV	GCCCAGGCCACGCCTC**T**CTGGAAAGTCCCCAGC
HIV 3’ LTR 377 SDM FWD-NEW	GGGGACTTTCCAGG**A**AGGCGTGGCCTGGGCGGG
HIV 3’ LTR 377 SDM REV-NEW	CCCGCCCAGGCCACGCCT**T**CCTGGAAAGTCCCC
HIV 3’ LTR 379 SDM FWD	GGGGACTTTCCAGGGA**A**GCGTGGCCTGGGCGGG
HIV 3’ LTR 379 SDM REV	CCCGCCCAGGCCACGC**T**TCCCTGGAAAGTCCCC
HIV 3’ LTR 384 SDM FWD	TTTCCAGGGAGGCGT**A**GCCTGGGCGGGACTG
HIV 3’ LTR 384 SDM REV	CAGTCCCGCCCAGGC**T**ACGCCTCCCTGGAAA
HIV 3’ LTR 389 SDM FWD	CAGGGAGGCGTGGCCT**A**GGCGGGACTGGGGAGT
HIV 3’ LTR 389 SDM REV	ACTCCCCAGTCCCGCC**T**AGGCCACGCCTCCCTG
HIV 3’ LTR 390 SDM FWD	AGGGAGGCGTGGCCTG**A**GCGGGACTGGGGAGTG
HIV 3’ LTR 390 SDM REV	CACTCCCCAGTCCCGC**T**CAGGCCACGCCTCCCT
HIV 3’ LTR 393 SDM FWD	GAGGCGTGGCCTGGGC**A**GGACTGGGGAGTGGCG
HIV 3’ LTR 393 SDM REV	CGCCACTCCCCAGTCC**T**GCCCAGGCCACGCCTC
HIV 3’ LTR 394 SDM FWD	GGCGTGGCCTGGGCG**A**GACTGGGGAGTGGCG
HIV 3’ LTR 394 SDM REV	CGCCACTCCCCAGTC**T**CGCCCAGGCCACGCC
HIV 3’ LTR 395 SDM FWD	GGCGTGGCCTGGGCGG**A**ACTGGGGAGTGGCGAG
HIV 3’ LTR 395 SDM REV	CTCGCCACTCCCCAGT**T**CCGCCCAGGCCACGCC
HIV 3’ LTR 399 SDM FWD	TGGCCTGGGCGGGACT**A**GGGAGTGGCGAGCCCT
HIV 3’ LTR 399 SDM REV	AGGGCTCGCCACTCCC**T**AGTCCCGCCCAGGCCA
HIV 3’ LTR 400 SDM FWD	GGCCTGGGCGGGACTG**A**GGAGTGGCGAGCCCTC
HIV 3’ LTR 400 SDM REV	GAGGGCTCGCCACTCC**T**CAGTCCCGCCCAGGCC
HIV 3’ LTR 401 SDM FWD	GCCTGGGCGGGACTGG**A**GAGTGGCGAGCCCTCA
HIV 3’ LTR 401 SDM REV	TGAGGGCTCGCCACTC**T**CCAGTCCCGCCCAGGC
HIV 3’ LTR 402 SDM FWD	CCTGGGCGGGACTGGG**A**AGTGGCGAGCCCTCAG
HIV 3’ LTR 402 SDM REV	CTGAGGGCTCGCCACT**T**CCCAGTCCCGCCCAGG

HIV 3’ LTR 406 SDM FWD	GGCGGGACTGGGGAGT**A**GCGAGCCCTCAGATGC
HIV 3’ LTR 406 SDM REV	GCATCTGAGGGCTCGC**T**ACTCCCCAGTCCCGCC
HIV 3’ LTR 409 SDM FWD	GGACTGGGGAGTGGC**A**AGCCCTCAGATGCTG
HIV 3’ LTR 409 SDM REV	CAGCATCTGAGGGCT**T**GCCACTCCCCAGTCC
HIV 3’ LTR 454 SDM FWD	GCTTTTTGCCTGTACT**A**GGTCTCTCTGGTTAGA
HIV 3’ LTR 454 SDM REV	TCTAACCAGAGAGACC**T**AGTACAGGCAAAAAGC
HIV 3’ LTR 464 SDM FWD	GTACTGGGTCTCTCT**A**GTTAGACCAGATCTG
HIV 3’ LTR 464 SDM REV	CAGATCTGGTCTAAC**T**AGAGAGACCCAGTAC
HIV 3’ LTR A3F SDM 474 FWD	CTCTCTGGTTAGACCA**A**ATCTGAGCCTGGGAGC
HIV 3’ LTR A3F SDM 474 REV	GCTCCCAGGCTCAGAT**T**TGGTCTAACCAGAGAG
HIV 3’ LTR 479 SDM FWD	GGTTAGACCAGATCT**A**AGCCTGGGAGCTCTC
HIV 3’ LTR 479 SDM REV	GAGAGCTCCCAGGCT**T**AGATCTGGTCTAACC
HIV 3’ LTR 485 SDM FWD	GACCAGATCTGAGCCT**A**GGAGCTCTCTGGCTAA
HIV 3’ LTR 485 SDM REV	TTAGCCAGAGAGCTCC**T**AGGCTCAGATCTGGTC
HIV 3’ LTR 496 SDM FWD	GCCTGGGAGCTCTCT**A**GCTAACTAGGGAACC
HIV 3’ LTR 496 SDM REV	GGTTCCCTAGTTAGC**T**AGAGAGCTCCCAGGC
HIV 3’ LTR 505 SDM FWD	GCTCTCTGGCTAACTA**A**GGAACCCACTGCTTAA
HIV 3’ LTR 505 SDM REV	TTAAGCAGTGGGTTCC**T**TAGTTAGCCAGAGAGC
HIV 3’ LTR A3F SDM 506 FWD	CTCTCTGGCTAACTAG**A**GAACCCACTGCTTAAG
HIV 3’ LTR A3F SDM 506 REV	CTTAAGCAGTGGGTTC**T**CTAGTTAGCCAGAGAG
HIV 3’ LTR 542 SDM FWD	AATAAAGCTTGCCTT**A**AGTGCTTCAAGTAGT
HIV 3’ LTR 542 SDM REV	ACTACTTGAAGCACT**T**AAGGCAAGCTTTATT
HIV 3’ LTR A3F SDM 577 FWD	TGCCCGTCTGTTGTGT**A**ACTCTGGTAACTAGAG
HIV 3’ LTR A3F SDM 577 REV	CTCTAGTTACCAGAGT**T**ACACAACAGACGGGCA
HIV 3’ LTR 583 SDM FWD	CTGTTGTGTGACTCT**A**GTAACTAGAGATCCC

HIV 3’ LTR 583 SDM REV	GGGATCTCTAGTTAC**T**AGAGTCACACAACAG
HIV 3’ LTR A3F SDM 593 FWD	GACTCTGGTAACTAGA**A**ATCCCTCAGACCCTTT
HIV 3’ LTR A3F SDM 593 REV	AAAGGGTCTGAGGGAT**T**TCTAGTTACCAGAGTC
HIV 3’ LTR A3F SDM 602 FWD	AACTAGAGATCCCTCA**A**ACCCTTTTAGTCAGTG
HIV 3’ LTR A3F SDM 602 REV	CACTGACTAAAAGGGT**T**TGAGGGATCTCTAGTT
HIV 3’ LTR SDM 365 revert FWD	GGGACTTTCCGCTGG**G**GACTTTCCAGGGAGG
HIV 3’ LTR SDM 365 revert REV	CCTCCCTGGAAAGTC**C**CCAGCGGAAAGTCCC
HIV 3’ LTR SDM 384 revert FWD	TTTCCAGGGAGGCGT**G**GCCTGGGCAGGACTA
HIV 3’ LTR SDM 384 revert REV	TAGTCCTGCCCAGGC**C**ACGCCTCCCTGGAAA
HIV 3’ LTR SDM 393 revert FWD	AGGCGTAGCCTGGGC**G**GGACTAATGAGTGGC
HIV 3’ LTR SDM 393 revert REV	GCCACTCATTAGTCC**C**GCCCAGGCTACGCCT
HIV 3’ LTR SDM 399 revert FWD	AGCCTGGGCAGGACT**G**ATGAGTGGCGAGCCC
HIV 3’ LTR SDM 399 revert REV	GGGCTCGCCACTCAT**C**AGTCCTGCCCAGGCT
HIV 3’ LTR SDM 400 revert FWD	GCCTGGGCAGGACTA**G**TGAGTGGCGAGCCCT
HIV 3’ LTR SDM 400 revert REV	AGGGCTCGCCACTCA**C**TAGTCCTGCCCAGGC
HIV 3’ LTR SDM 464 revert FWD	GTACTGGGTCTCTCT**G**GTTAGACCAGATCTA
HIV 3’ LTR SDM 464 revert REV	TAGATCTGGTCTAAC**C**AGAGAGACCCAGTAC
HIV 3’ LTR SDM 485 revert FWD	ACCAGATCTAAGCCT**G**GGAGCTCTCTAGCTA
HIV 3’ LTR SDM 485 revert REV	TAGCTAGAGAGCTCC**C**AGGCTTAGATCTGGT
HIV 3’ LTR SDM 542 revert FWD	AATAAAGCTTGCCTT**G**AGTGCTCAAAGTAGT
HIV 3’ LTR SDM 542 revert REV	ACTACTTTGAGCACT**C**AAGGCAAGCTTTATT
HIV 3’ LTR SDM 583 revert FWD	CTGTTGTGTGACTCT**G**GTAACTAGAGATCCC
HIV 3’ LTR SDM 583 revert REV	GGGATCTCTAGTTAC**C**AGAGTCACACAACAG

* Note to Langlois Lab: The numbers used in this thesis are in conventional orientation. In order to match the numbers used within the laboratory, subtract these numbers from 635 (ex. SDM 485 in this thesis would be 635 - 485 = SDM 150 in the lab, and vice versa).

8.2 Cloning Strategies

Figure A1. Cloning Strategy for LTR-eGFP Reporter Constructs.

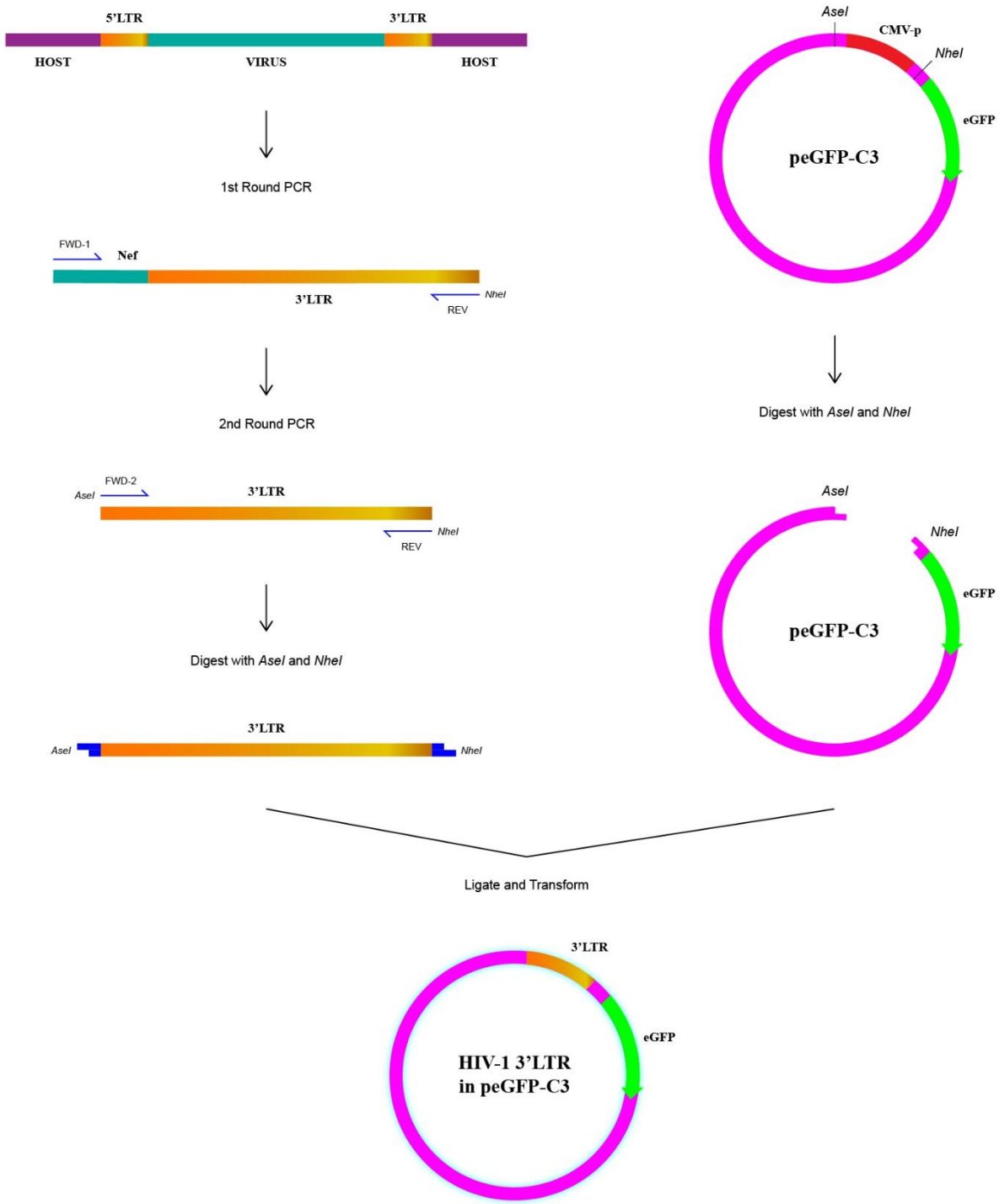


Figure A2. Cloning Strategy for Reconstituted Heavily Mutated Clones.

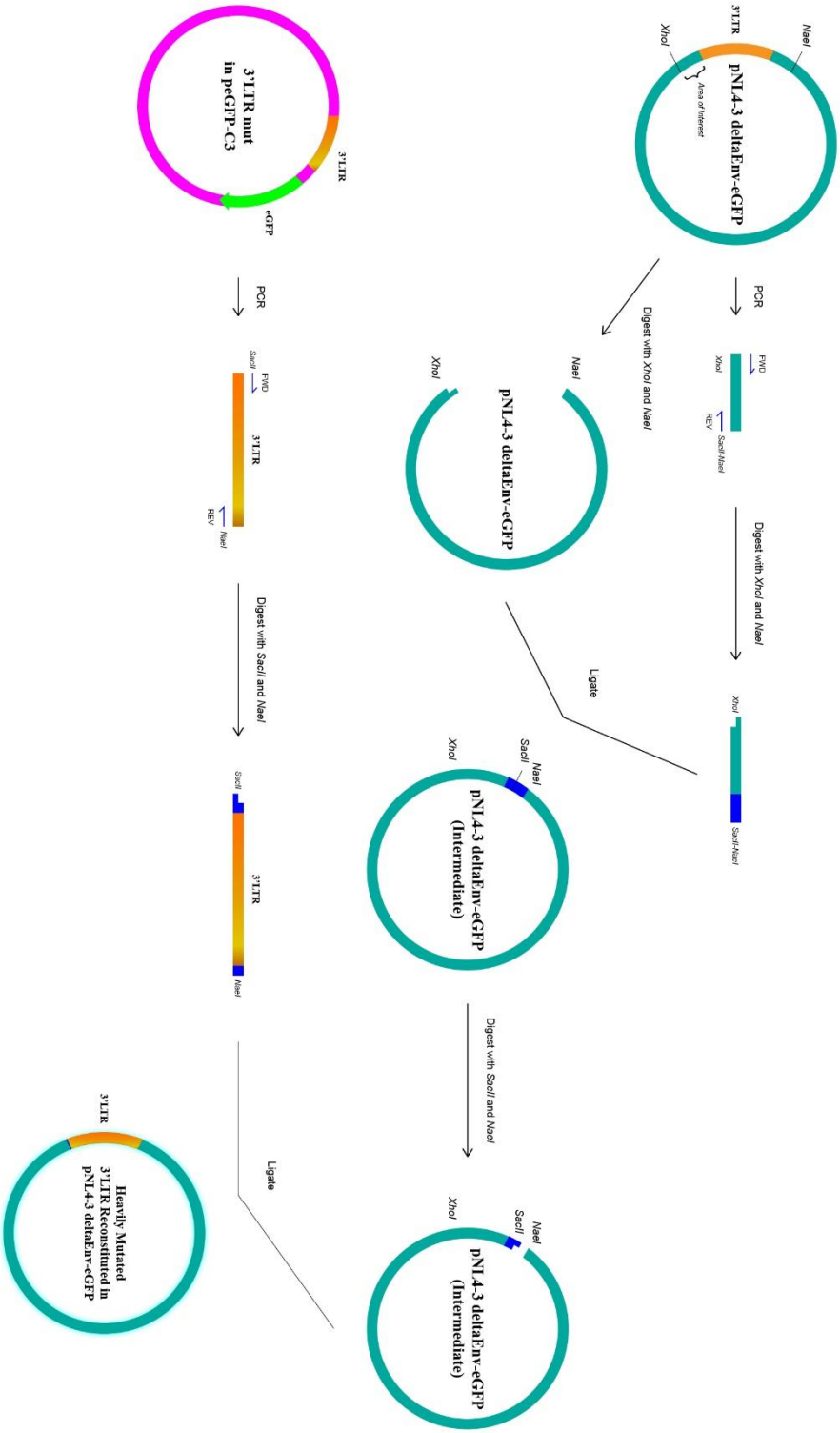


Figure A3. Cloning Strategy for Reconstituted Revertant Clones.

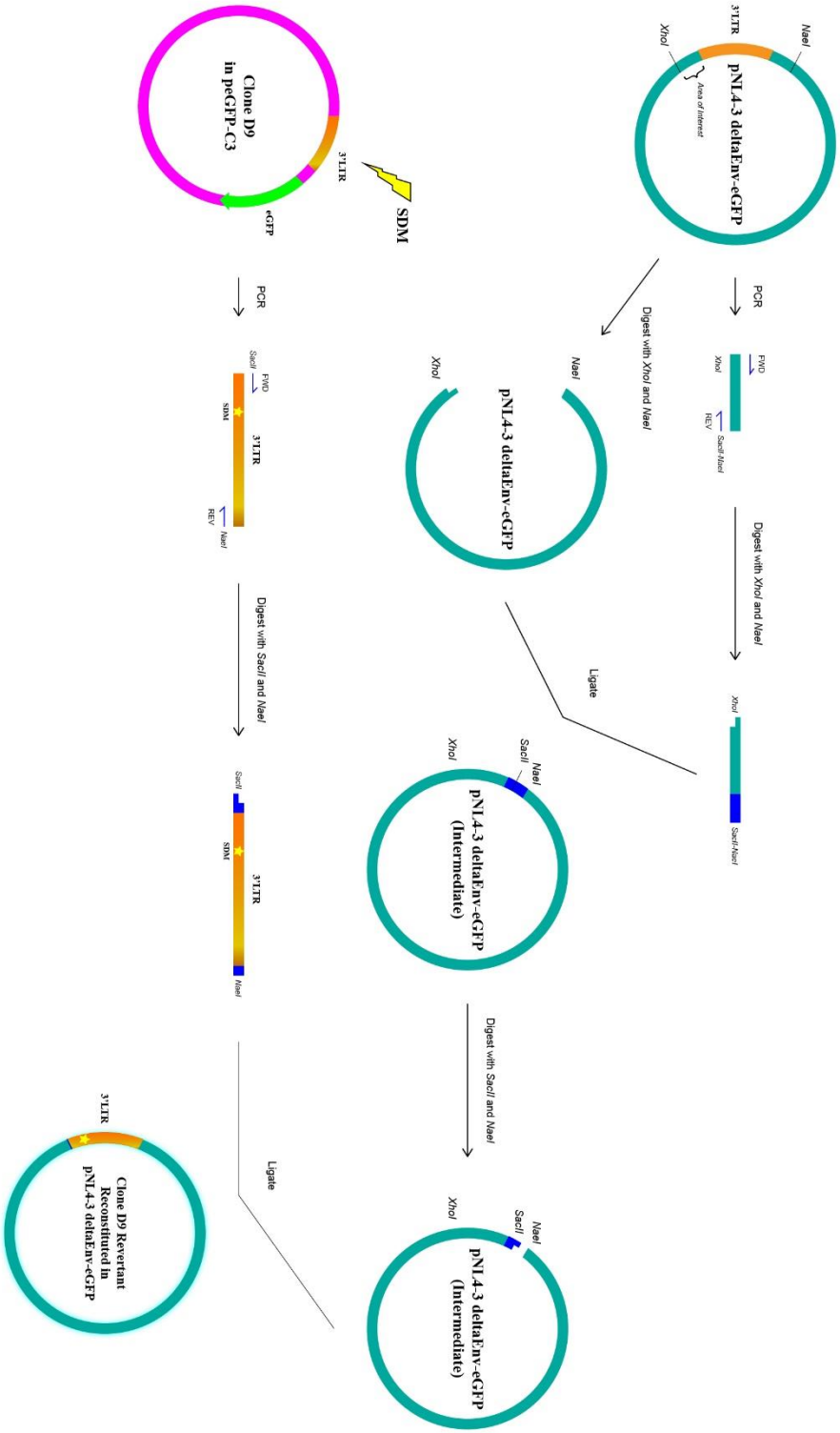


Figure A4. Cloning Strategy for Reconstituted Single/Double Point Mutants.

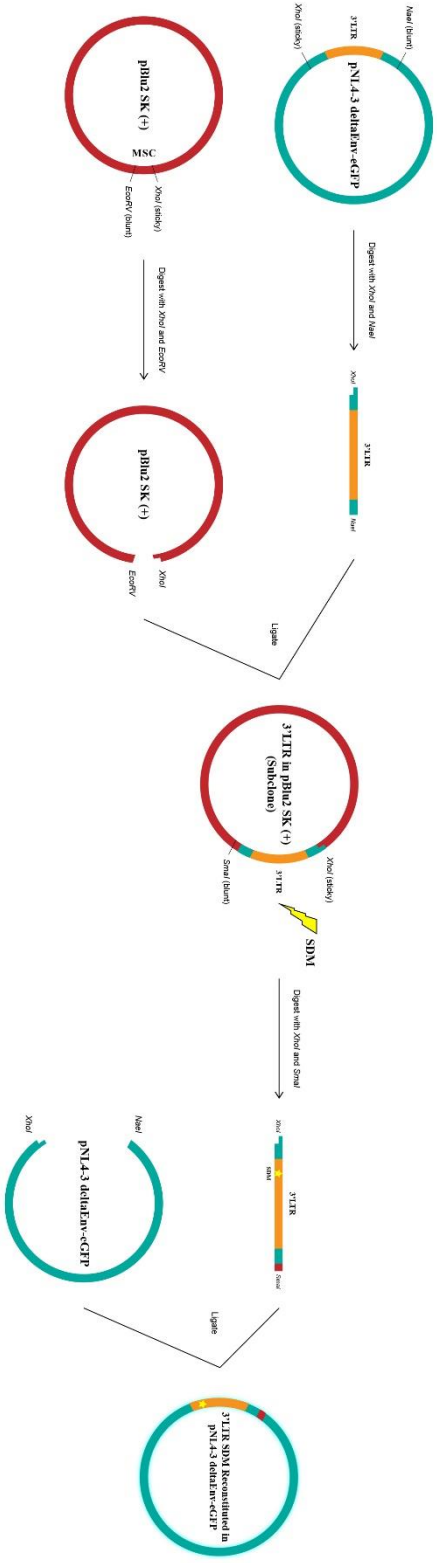
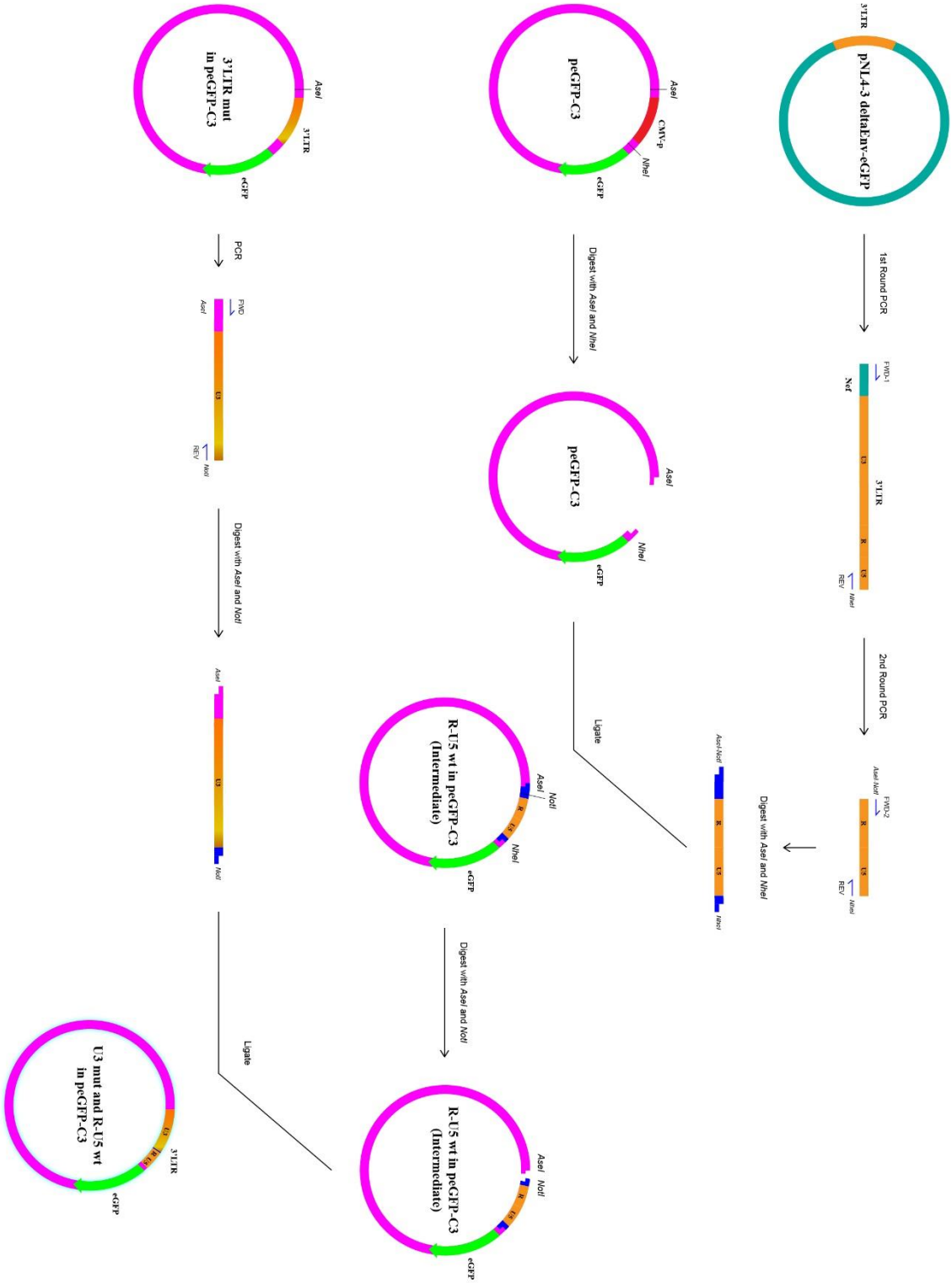


Figure A5. Cloning Strategy for U3 mut/R-U5 wt Clones.



8.3 Reprint Permissions

Reference (37) for Figure 1.3.

Permission is granted for the use of Fig 1, p.3 in the article detailed below in your Masters thesis only. Please include complete reference and copyright to Cold Spring Harbor Laboratory Press.

Best wishes for success with your thesis,

Carol C Brown
Permissions Coordinator
brown@csh.edu

Name: Cindy Lam
Company/Institution: University of Ottawa
Library Address: 451 Smyth Road
Library Address (line 2):
City: Ottawa
State (US and Canada): Ontario
Country: Canada
Zip: K1H8L1
Title: Master's student
Lab/Department: Laboratory of Dr. Marc-André Langlois / Department of Biochemistry, Microbiology and Immunology
Phone: (613) 562-5800 ext. 7111
Fax: (613) 562-5452
Email: *omitted for privacy*
Title of Publication: Investigation of the Contribution of APOBEC3-Induced Mutations in HIV-1 Latency (tentative)
Authors/Editors: Cindy Lam
Date of Publication: August 2019 (tentative)
Publisher: University of Ottawa
Title of CSHLP Journal/Book: Cold Spring Harbor Perspectives Medicine
Title of Article/Chapter: HIV-1 Reverse Transcription
CSHL Authors/Editors: Wei-Shau Hu, Stephen H. Hughes
Page Numbers: 3
Figure Numbers: 1
Figure Page Numbers: 3
Copyright Date: 2012
Language: English
Territory:
Format:
Additional comments:

Reference (46) for Figure 1.4.

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

This Agreement between Miss. Cindy Lam ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4606660569226
License date	Jun 12, 2019
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Genetics
Licensed Content Title	The causes and consequences of HIV evolution
Licensed Content Author	Andrew Rambaut et al
Licensed Content Date	Jan 1, 2004
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	no
Title	Investigation of the Contribution of APOBEC3-Induced Mutations in HIV-1 Latency
Institution name	University of Ottawa
Expected presentation date	Jun 2019
Portions	Figure 1: Key aspects of the HIV life cycle.
Requestor Location	*omitted for privacy*
Total	0.00 USD

Terms and Conditions

Springer Nature Terms and Conditions for RightsLink Permissions

Springer Nature Customer Service Centre GmbH (the Licensor) hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and

requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).
   If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.
2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.
3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.
4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.
5. Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).
6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.
7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.
8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.
9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.
10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

Appendix — Acknowledgements:

For Journal Content:
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

For Adaptations/Translations:
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT]** (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

For Book content:
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

Other Conditions:

Version 1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Reference (41) for Figure 1.5.

ELSEVIER LICENSE
TERMS AND CONDITIONS

This Agreement between Miss. Cindy Lam ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4606630683448
License date	Jun 12, 2019
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell
Licensed Content Title	An Integrated Overview of HIV-1 Latency
Licensed Content Author	Debbie S. Ruelas,Warner C. Greene
Licensed Content Date	Oct 24, 2013
Licensed Content Volume	155
Licensed Content Issue	3
Licensed Content Pages	11
Start Page	519
End Page	529
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 2. Major Mechanisms of HIV-1 Latency in CD4 T cells
Title of your thesis/dissertation	Investigation of the Contribution of APOBEC3-Induced Mutations in HIV-1 Latency
Expected completion date	Jun 2019
Estimated size (number of pages)	100
Requestor Location	*omitted for privacy*
Publisher Tax ID	GB 494 6272 12
Total	0.00 USD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as

well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:
15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use
    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all

value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.
Policies for sharing publishing journal articles differ for subscription and gold open access articles:
Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.
Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.
If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.
Please refer to Elsevier's posting policy for further information.
18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.
19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions
You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.
Terms & Conditions applicable to all Open Access articles published with Elsevier:
Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.
The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.
Additional Terms & Conditions applicable to each Creative Commons user license:
CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.
CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.
CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.
Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.9

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Reference (124) for Figure 1.6.

ELSEVIER LICENSE
TERMS AND CONDITIONS

This Agreement between Miss. Cindy Lam ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4606630237847
License date	Jun 12, 2019
Licensed Content Publisher	Elsevier
Licensed Content Publication	Virology
Licensed Content Title	APOBECs and virus restriction
Licensed Content Author	Reuben S. Harris,Jaquelin P. Dudley
Licensed Content Date	May 1, 2015
Licensed Content Volume	479
Licensed Content Issue	n/a
Licensed Content Pages	15
Start Page	131
End Page	145
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Fig. 2. Model for HIV-1 restriction by APOBEC3 enzymes and virus protection by Vif.
Title of your thesis/dissertation	Investigation of the Contribution of APOBEC3-Induced Mutations in HIV-1 Latency
Expected completion date	Jun 2019
Estimated size (number of pages)	100
Requestor Location	*omitted for privacy*
Publisher Tax ID	GB 494 6272 12
Total	0.00 USD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as

well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:
15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use
    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all

value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.
Policies for sharing publishing journal articles differ for subscription and gold open access articles:
Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.
Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.
If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.
Please refer to Elsevier's posting policy for further information.
18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.
19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions
You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.
Terms & Conditions applicable to all Open Access articles published with Elsevier:
Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.
The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.
Additional Terms & Conditions applicable to each Creative Commons user license:
CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.
CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.
CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.
Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.9

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.