Characterization of APOBEC3 Repertoires and Effects on Retroviral Replication

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in the
Department of Biochemistry, Microbiology, and Immunology
University of Saskatchewan
Saskatoon, SK
By
Tyson Follack

Permission to Use

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

Or

Dean's Office
College of Medicine
5D40 Health Science Building, 107 Wiggins Road
Saskatoon, Saskatchewan S7N 5E5
Canada

Disclaimer

The [name of company/corporation/brand name and website] were exclusively created to meet the thesis and/or exhibition requirements for the degree of [title of degree] at the University of Saskatchewan. Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Biochemistry,

Microbiology, and Immunology

University of Saskatchewan

Health Science Building

107 Wiggins Road

Saskatoon, Saskatchewan, Canada S7N 5E5

Abstract

APOBEC3 (A3) enzymes are a family of intrinsic retroviral restriction factors that are coordinately expressed in CD4+ T-cells and function to restrict retroviral replication. In the case of HIV-1, this primarily occurs in the absence of the HIV-1 protein, viral infectivity factor (Vif). Vif induces the polyubiquitination and degradation of A3 enzymes using the host proteasome pathway. The A3 enzyme family are single stranded DNA deaminases, capable of deaminating cytosine to form promutagenic uracil on single-stranded DNA substrates. In humans there are seven paralogs that comprise the A3 family, including: A3A, A3B, A3C, A3D, A3F, A3G, and A3H. However, only four paralogs have been associated with inhibition of retroviral replication in a large majority of the human populations, including: A3D, A3F, A3G and A3H. These four key enzymes have well-characterized repertoires of single nucleotide polymorphisms (SNPs) that affect both cellular stability and deaminase activity. Each repertoire of polymorphisms is variable in human populations and is influenced by human ancestry. Here we have undertaken a Saskatchewan-based mixed population study examining human genotypes for common APOBEC3H (A3H), APOBEC3F (A3F) and APOBEC3G (A3G) SNPs and how combinatory SNP variations affect HIV-1 replication. Using Sanger-sequencing based genotyping we were able to isolate dominant polymorphisms in our population. We found that the dominant genotypes in our population were heterozygous for A3F 231V and 231I SNPs, homozygous for the A3G 186H SNP, and predominantly inactive A3H SNP profiles. We tested the dominant A3F and A3G polymorphisms for cellular stability, viral packaging and restriction capacity, and found that the A3F 231V SNP provided the most robust restriction response when co-expressed with both A3G and A3F 231I. We also show that A3F 231V can hetero-oligomerize with both A3G and A3F 231I polymorphic variants, resulting in greater cellular stabilization and steady-state cellular expression in the presence and absence of HIV-1 Vif. The observed rise in cellular stability and virion packaging when A3F 231V was co-expressed with A3F 231 and A3G had a positive correlation with HIV-1 replication restriction efficiency. The various states of oligomerization between co-expressed A3 enzymes demonstrates that whole genotype analysis of A3 repertoires is essential in accurately understanding host-pathogen interactions on a population level.

Acknowledgements

I would first wish to thank Dr. Linda Chelico for accepting me into her lab in both undergraduate and graduate settings. Linda provides unparalleled support for all of her students, with hands-on training, daily guidance, as well as both scientific and emotional support. I will forever appreciate the time spent in her lab and lessons learned along the way.

I appreciate the guidance and direction provided by my advisory committee. Dr. Wei Xiao and Dr. Sylvia van den Hurk supplied constructive feedback for my project progression through the years, ultimately shaping the best possible scientific output.

The entirety of the Chelico Lab has provided the very definition of an ideal work environment and it is my fellow lab members I would like to thank for creating this dynamic. Dr. Madison Adolf, Dr. Cate Feng, Robin Love, Nazanin Mohammadzadeh, Lai Wong and Mo Khalil have individually contributed to a net of comradery that made everyday worth while.

I would also like to personally thank Dr. Steve Sanche and Dr. Kris Stewart for help in the acquisition of our first sets of blood donations in the lab, as well as all volunteers that donated for our cause. Without these generous donations this project would not be possible and for that I am grateful.

Funding is an essential component to research and I would like to thank the Canadian Institute for Health Research (CIHR) for project grant funding, as well as, National Science and Engineering Research Council (NSERC)-CREATE and College of Medicine funding sources for making it possible for me to pursue graduate studies.

Finally, I would like to thank and acknowledge my parents, Blaine and Coreen Follack, as well as my brothers Justin and Jordin Follack. They have forever been my rock through tough times and my scientific career was of no exception. They are always just a phone call away and have been the keystone to my success.

Table of Contents

Table of Figures

List of Tables

List of Abbreviations

nt	Nucleotide (s)
(-) DNA	minus strand/template strand
(+) DNA	positive strand/coding strand
3'UTR	3' untranslated region
A3	APOBEC3
A3H hap I	APOBEC3H haplotype I
A3H hap II	APOBEC3H haplotype II
A3H hap III	APOBEC3H haplotype III
A3H hap IV	APOBEC3H haplotype IV
A3H hap V	APOBEC3H haplotype V
A3H hap V	APOBEC3H haplotype VI
A3H hap VII	APOBEC3H haplotype VII
AID	Activation Induced Cytosine Deaminase
AIDS	Acquired Immunodeficiency Syndrome
APE	apurinic/apryidinic endonuclease
APOBEC1	apolipoprotein B mRNA editing, catalytic-enzyme polypeptide like 1
APOBEC2	apolipoprotein B mRNA editing, catalytic-enzyme polypeptide like 2
APOBEC3	apolipoprotein B mRNA editing, catalytic-enzyme polypeptide like 3
BER	base excision repair
bp	Base pair
Co-IP	Coimmunoprecipitation
CRCL5	Culling RING ubiquitin ligase 5
CTD	C-terminal domain
Cul5	Cullin 5
dNTP	deoxyribonucleoside triphosphate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EloB	Elongin B
EloC	Elongin C

Env	Envelope
FAM	Fluorescein
Gag	Group specific antigen
HIV-1	Human Immunodeficiency Virus Type 1
HIV-2	Human Immunodeficiency Virus Type 2
LTR	Long terminal repeat
mRNA	Messenger RNA
NC	Nucleocapsid
PBMC	Peripheral blood mononuclear cell
Pol	polymerase
PPT	Poly-purine tract
PR	Protease
RBX2	Ring box protein 2
RNase A	Ribonuclease A
RNase H	Ribonuclease H
RPA	Replication protein A
RT	Reverse transcriptase
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIV	Simian Immunodeficiency Virus
SIVcpz	Simian immunodeficiency virus from chimpanzees
SNP	Single nucleotide polymorphism
SOCS2	Suppressor of cytokine signaling 2
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
UDG/UNG	uracil DNA glycosylase
Vif	Viral Infectivity Factor

1.0 Literature Review

1.1 Evolution and diversification of APOBEC enzymes

The apolipoprotein B mRNA editing, catalytic-enzyme polypeptide like 3 (APOBEC3) enzyme family are single-stranded (ss) cytidine deaminases, capable of deaminating cytosine motifs in both lentiviral and host ssDNA (*1-3*). Some A3s can deaminate single-stranded RNA (ssRNA), but the function of this activity is not known (*4*). Derivatives of APOBEC enzymes are found not only in placental mammals, but also in ancestors of jawless-fish dating back over 500 million years. Activation induced deaminase (AID) is the most ancestral variant (Figure 1.1) (*5*). AID is an ssDNA cytidine deaminase that is expressed in B-cells and is responsible for antibody maturation by initiating class-switch recombination and somatic hypermutation (*5*). As jawed vertebrates evolved, duplications of AID produced APOBEC1 (A1), APOBEC2 (A2) and APOBEC (A3) members of the APOBEC superfamily (*5*). Although A1 and A2 evolved prior to A3, the physiological function of A1 appears to only involve RNA editing, while A2 demonstrates no deaminase activity (*6, 7*). Until recently, the apex of APOBEC gene duplication was thought to be in the genomes of both humans and non-human primates, whereby a cluster of seven APOBEC3 genes are encoded solely on chromosome 22. Other mammalian relatives only encode up to four APOBEC3 genes (*8-10*). However, bats have been found to encode at least 20 different APOBEC3 genes (*11*) . In humans, each A3 loci is also variable, with prevalent polymorphic regions that vary in human populations, suggesting that A3 repertoires have yet to stabilize in placental mammals (*12-14*). The selective pressures influencing A3 evolution are largely unknown, but historical modulation in retroviral challenge has been theorized to alter the requirement for defense proteins, like A3 deaminases (*6, 14, 15*).

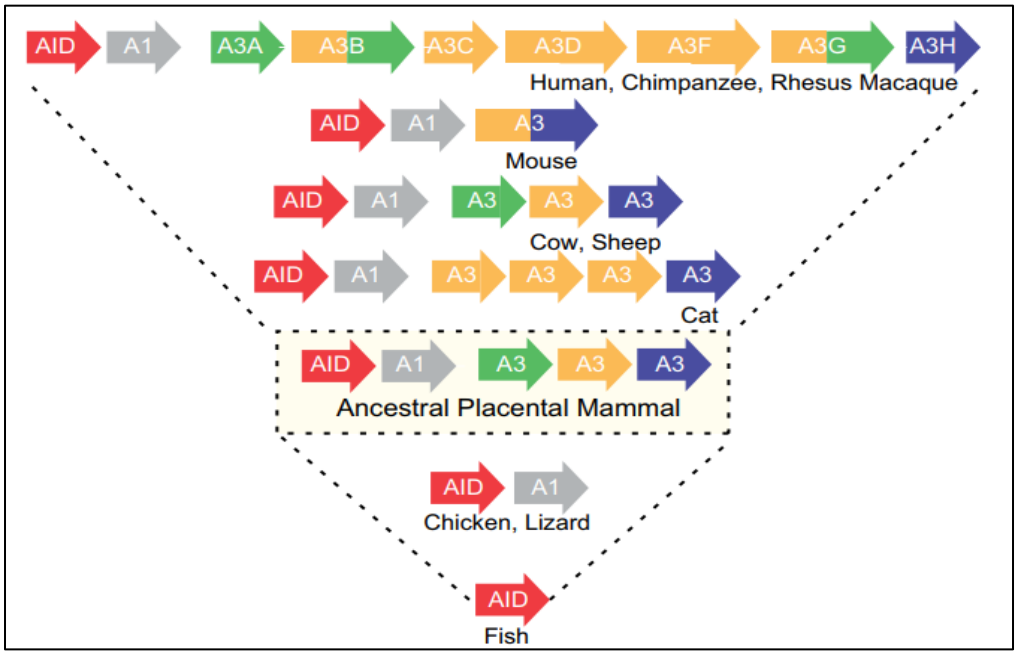


Figure 1. 1 Phylogenetic origins of APOBEC family members.

The family originated with Activation Induced Deaminase (AID) in jawless fish. Timing of gene duplication events leading to the A3 subgroup is unclear, but likely occurred around 200 million years ago. Mammals express the largest diversity of A3 enzymes. Reprinted with permission from (16).

1.2 Mechanism of APOBEC-mediated cytosine deamination and repair

Mammalian APOBEC3 deaminases scan ssDNA substrates and catalytically act on specific cytosine-containing motifs, deaminating cytosine to from uracil (Figure 1.2, 1.3) (*1-3, 17*). Because uracil does not have a coding function in DNA it is defined a pro-mutagenic lesion. If Uracil-based DNA lesions are promptly detected they can repaired by host cell DNA repair pathways, if not, they can be used as an erroneous template for replication and result in mutations *(1-3)*.

Base Excision Repair (BER) is a DNA repair mechanism initiated by a lesion specific repair enzyme. Specific to uracils, the main enzyme is a member of the Uracil DNA Glycosylase (UDG) superfamily; Uracil DNA N-Glycosylase (UNG), which catalyzes the cleavage of the N-glycosylic bonds between the uracil containing substrate and its deoxyribose backbone, creating an abasic site (*18-20*). The abasic site is then incised by an AP endonuclease on the 5′ -side of the abasic site, creating a gap with a 3′ -OH overhang for later polymerase priming on a single strand. The most prominent AP endonuclease in mammals being APE1, which accounts for over 95% of cellular endonuclease activity (*21*). Using the 3′ -OH over-hang left by the APE1, a DNA polymerase will synthesize DNA to fill the gap. Under normal circumstances the result is restoration of the original nucleotide sequence due to the action of a high fidelity polymerase (*9*). If an error prone polymerase is recruited to the abasic site due to cell stress, then any base can be inserted into the DNA. Double stranded (ds) DNA breaks can also arise, but requires uracils on both strands in close proximity, and are repaired using host non-homologous end-joining (*22*).

An alternative fate to uracil lesion repair is the faulty recognition of uracil as thymine during DNA replication. That is, during DNA replication if uracil is recognised as thymine opposing strand synthesis will add an adenine across from uracil, resulting in a uracil-adenine intermediate. Through BER, the U-A intermediate is converted to a thymine-adenine base pair. The result being the transition mutation from a Cytosine-Guanine (C-G) base pair prior to A3 deamination to a Thymine-Adenine (T-A) base pair (*23*).

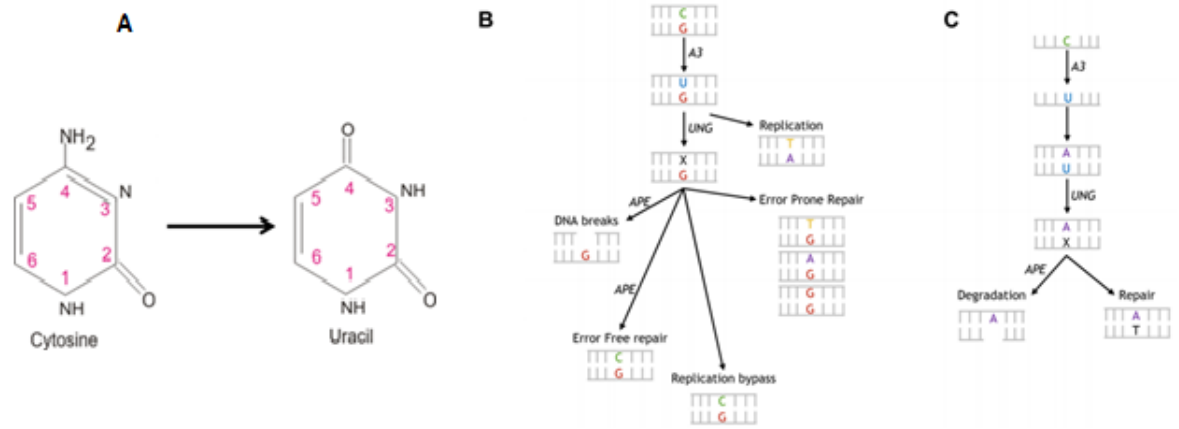


Figure 1. 2 Cytosine deamination and uracil repair

(A) Structural similarities between thymine and uracil result in potential transition mutations following cytosine deamination events. (B) Diagram indicating the fate of uracil after cytosine deamination in genomic DNA, which results after the activity of AID. If uracil-based lesions fail to be repaired, the uracil is replaced with thymine, resulting in a T/A base pair instead of a C/G base pair. Alternatively, if uracils are created by cytosine deamination they are removed by uracil DNA glycosylase (UNG) to create an abasic site. APE processes abasic sites by cleaving the DNA backbone at the respective site. The cleavage reaction prompts host DNA repair pathways to repair the abasic site. DNA breaks can also occur. (C) A3 enzymes deaminate single-stranded (−) DNA substrates, which are present as replication intermediates in retroviruses. Uracils formed by cytosine deamination in the (−) DNA strand serves as a template for the reverse transcriptase during (+) DNA synthesis, resulting in G → A mutation in the (+) DNA. Depending on the number of uracils, this can lead to degradation of the template in the nucleus or repair of uracil and integration of proviral DNA. Reprinted with permission from (*4*).

1.3 Effects of APOBEC3 deamination on different ssDNA substrates

However unlikely it may seem, either result of A3 mediated deamination correction or mutation have both costs and benefits to cellular fitness. The scales tip depending on the source of available ssDNA substrate. That is, if ssDNA is presented as a transposing retroelement or a ssDNA intermediate of an invading retrovirus, A3 enzymes will promote mutagenesis as a defence mechanism to block transposition or replication respectively; referred to as deamination dependent restriction (*1, 24-26*). A3-mediated deamination dependent restriction can result in the functional inactivation of essential genes, ultimately leading to DNA degradation by the host cell nucleases and re-purposed as cellular nutrients (*1, 24, 27*). Whereas, A3-mediated deamination independent mechanisms of replication restriction act cooperatively with deamination dependent pathways to strengthen the restriction response and do so without catalytic cytosine deamination (*25, 28, 29*). Rather, this mechanism inhibits second strand DNA synthesis of the virus by blocking the viral mechanism for reverse transcription. Reverse transcription can be inhibited by A3s in a 'roadblock' fashion, where A3s bound to ssDNA block the progression of viral reverse transcriptase (RT), which allows ssDNA intermediates to be susceptible to deamination for a longer duration of time (*30-34*).

However, due to the A3 enzyme's inability to distinguish between sources of ssDNA substrate, they also pose a risk for self-mutation of host genomic DNA during replication. That is, if ssDNA substrate is presented as host-cell genomic DNA during replication, A3 enzymes that can enter the nucleus can impose the same deaminations observed on foreign retroviral ssDNA (*21, 35, 36*). This trend is also exacerbated under conditions of replicative stress or when the cell cycle is altered, leaving ssDNA exposed for greater durations; events that are synonymous with human cancers (*37, 38*). Under these conditions, A3 specific mutations begin to rise and fuel mutagenic evolution of human cancers (*39-41*). This scenario indicates the necessity for cooperative regulatory mechanisms between nuclear A3 enzymes and DNA replication to ensure stability of the host genome.

1.4 Variation in functional characteristics between APOBEC3 paralogs

The A3 family is a highly diverse group of deaminases present in all placental mammals, and humans are of no exception. With seven characterized human A3 enzymes: A3A, A3B, A3C, A3D, A3F, A3G and A3H; although variable, all have been characterized to have catalytic activity on ssDNA substrates (*2, 4, 29*). Cellular localization plays a pivotal role in determining A3 enzyme function. Paralogs that have been observed to be localized to the nucleus: A3A, A3B, and A3H haplotype I, have all been associated with fueling mutagenesis in human cancers (*36, 39-44*). While A3s that have been associated with restriction of invading retroviruses: A3D, A3F, A3G, and A3H haplotypes II, V, VII are all localized to the cytoplasm (*33, 45-48*).

The cytoplasmic fraction of A3 paralogs can be further distinguished into a hierarchy based on their ssDNA deamination efficiency (*25, 26, 28, 29, 45, 46, 49-54*). Causative components for the observed variations have been described to include: oligomerization, RNA binding affinities, and ssDNA scanning mechanisms (*1, 50, 55-57*). A3G, A3F, and A3H are the most potent restrictors of retroviral replication (*50, 51, 58*). One of the most prominent characteristics of A3 enzymes that alters deaminase activity is ssDNA scanning (Figure 1.3). The mechanisms of scanning ssDNA include sliding, jumping or translocation by intersegmental transfer on ssDNA substrates (*50*). The ability to slide allows the enzyme to scan a local strand of ssDNA within 20 nucleotides (nt), jumping allows the enzyme to bypass short dsDNA segments or double stranded RNA-DNA hybrids on the same strand, and intersegmental transfer is similar to jumping, but requires the enzyme to simultaneously bind two ssDNA strands prior to strand transfer (*50*). The ssDNA scanning mechanisms vary between the A3 family and significantly affect the enzymes' ability to deaminate ssDNA substrates and induce mutagenesis (*50, 51, 58*). A3G and A3H can scan and translocate three dimensionally, with A3G being able to jump, and A3H being able to jump and undergo intersegmental transfer (*51, 58*). Conversely, A3F is unable to slide on ssDNA and can only jump, resulting in a processivity deficit of nearly 4-fold in comparison to A3G and A3H (*51, 52*). A3D is the least characterized from the A3 family; having minimal effect on retroviral restriction has resulted in less characterization of A3D's deamination and scanning mechanisms that have been greatly elucidated in other A3 paralogs (*45, 46, 53*).

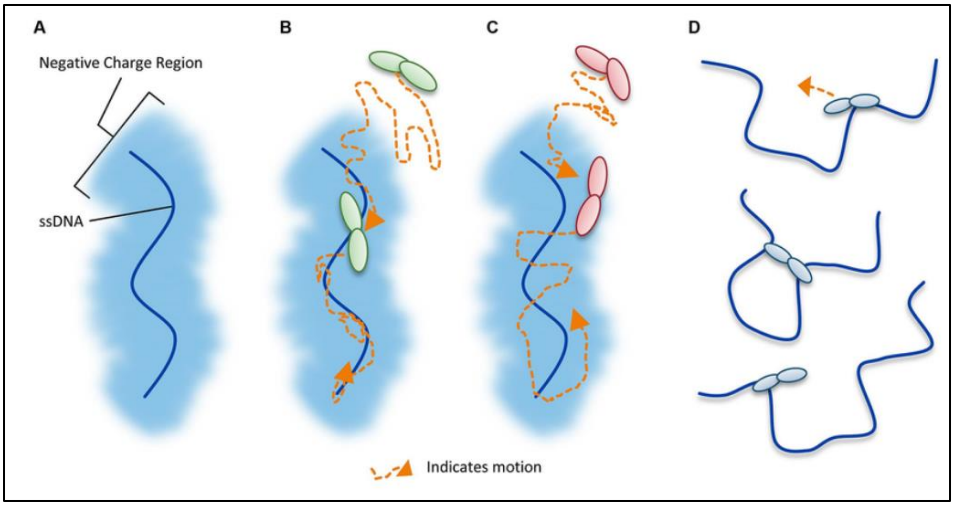


Figure 1. 3 APOBEC3 ssDNA scanning mechanisms

DNA scanning by facilitated diffusion. (A) Illustration of DNA showing the negatively charged region of DNA. (B) Illustration showing 1-dimensional DNA scanning by sliding. Dotted orange line indicates enzyme path. Sliding enables a thorough search of local areas of ssDNA. (C) Illustration depicting 3-dimensional DNA scanning by jumping. Jumping enables larger translocations on DNA substrates but lacks a local search process. Jumping dissociations from the DNA occur when the enzyme does not leave the negatively charged domain of the DNA. The enzyme then has a higher likelihood of rejoining with the same strand of DNA compared to dissociation into the bulk solution. (D) Illustration depicting 3-dimensional DNA scanning by intersegmental transfer. Intersegmental transfer enables larger translocations on DNA substrates but also lacks a local search process. An enzyme with two DNA-binding domains binds two regions of DNA simultaneously before dissociating from one region to move to another. (B–D) Enzyme in illustration is shown as a dimer, although the oligomerization states may vary with different A3 enzymes. Reprinted with permission from (*1*).

1.5 Restriction of Human Immunodeficiency Virus Type 1 by APOBEC3 deaminases

Evidence of APOBEC3 deaminase activity has been recorded in many fields, ranging from epigenetic gene regulation to human cancer studies, but the most characterized response is that of Human Immunodeficiency Virus Type 1 (HIV-1) replication restriction in human populations (*2, 3, 9*). Global prevalence and severity of HIV-1infections and AIDs progression has fueled studies in the characterization of HIV restriction factors like members of the A3 family, given the current absence of a cure for the disease (*59, 60*). HIV-1 is a retrovirus, which means that it has a single-stranded RNA genome and replicates using a ssDNA intermediate synthesized via reverse transcription (*61*). The nascent ssDNA is then used as a template for second strand or (+) strand DNA synthesis to form dsDNA, which can then undergo nuclear import and integration into the host chromosome. The RNA genome is protected by bound p6 nucleocapsid that is cleaved from the Gag polyprotein after virion maturation (*61*). The HIV-1 genome and accessory proteins are surrounded in a protective capsid, which is subsequently surrounded by a protein matrix and envelope derived from the cost cell plasma membrane (*61*). The gp120 and gp41 proteins are HIV-1 surface subunit and transmembrane proteins, respectively, that are on the surface of the virus outer membrane and facilitate viral entry and membrane fusion via binding to CD4 and CCR5/CXCR4 receptors on the surface of $CD4^+$ lymphocytes (*61*). Once bound, the fusogenic properties of the HIV-1 gp41 peptide facilitates the release of the HIV-1 capsid into the cell. The only host component required for reverse transcription and copying of the HIV-1 RNA genome into dsDNA is a pool of deoxynucleotide triphosphates (dNTPs), which diffuse into the capsid from the host cell cytoplasm, allowing the virus to synthesize the proviral DNA under the protection of its capsid proteins (*62, 63*). The HIV-1 capsid is only semipermeable, allowing small charged particles to diffuse through capsid pores, but excluding larger proteins, and A3 enzymes are of no exception. Thus, due to the lack of access to nascent ssDNA intermediates of the reverse transcribing virus, the catalytic activity of A3 proteins is evaded in the initial virus producer cell. A3 proteins overcome this barrier through promiscuous binding of viral RNA to enable encapsidation in newly assembled viral particles by binding RNAs which are also bound to HIV-1 Gag (*57, 64*). With A3 proteins now encased within the HIV-1 capsid, upon entry into a new target cell and initiation of reverse transcription, HIV-1 ssDNA becomes available for A3 mediated deamination to restrict replication of the invading retrovirus (Figure 1.4) (*16*).

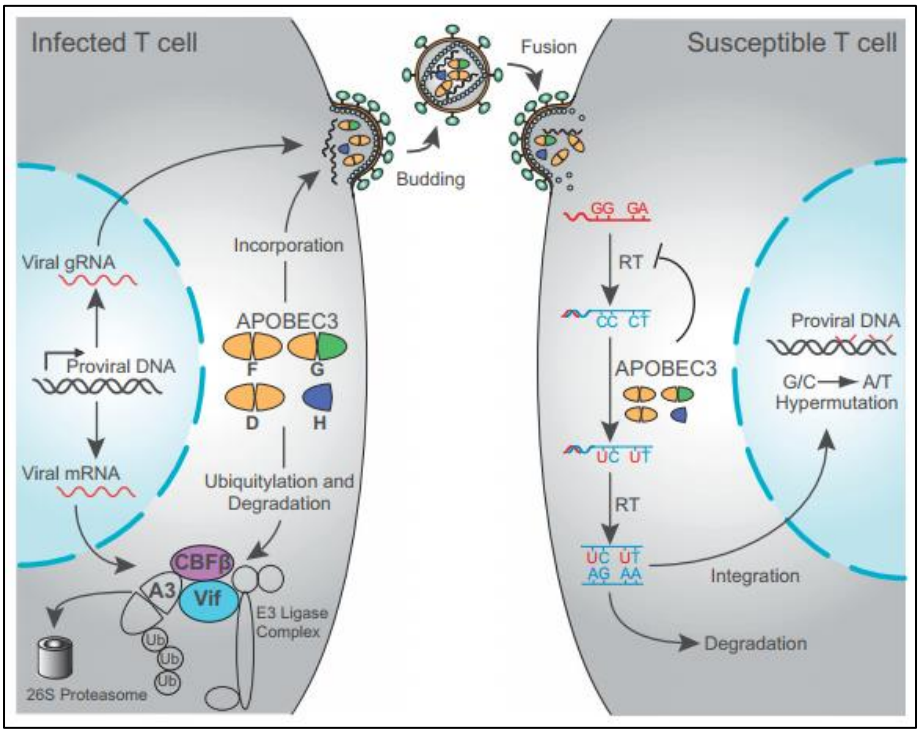


Figure 1. 4 Model for APOBEC3 mediated restriction of HIV-1

Model for HIV-1 restriction by APOBEC3 proteins. The figure depicts the necessary encapsidation of A3 proteins in producer CD4+ T-cells, which is facilitated by RNA binding domains associating with the RNA genome of the HIV-1 virus. This model depicts A3 deamination and RT-inhibition occurring in the target cell cytoplasm, however, this process occurs within the virus capsid. The A3s deaminate nascent ssDNA upon reverse transcription of (-) DNA. This results in degradation of viral DNA or integration of hypermutated proviral DNA (16). Each deamination event produces a uracil in place of a cytosine at specific motifs. Because HIV-1 lacks genome repair mechanisms the uracil base lesions remain and upon second strand synthesis, the uracil being recognized as a thymine causes an adenine to be inserted. Once integrated into the host genome, the U-A is repaired to T-A forming G-C transition mutations. This restriction activity is largely thwarted by the action of the viral protein Vif that recruits an E3 ligase complex and an A3 resulting in proteasomal degradation of A3s in the virus producer cell. Reprinted with permission from (16).

1.6 Inactivation of APOBEC3 deaminases by HIV-1 Vif

HIV-1 has an inherent ability to adapt and escape the host immune response due to high mutational frequencies observed throughout the course of infection (*65*). The frequency of genetic variation is due to collaborative efforts of the HIV-1 RT that lacks proof reading, recombination between co-packaged HIV-1 genomes, and positive selective pressures from the host (*66-69*). These mechanisms of adaption have also allowed HIV-1 to develop defences against A3 deaminases through the action of the HIV-1 Virus Infectivity Factor (Vif).

Vif is an HIV-1 accessory protein that binds to and targets A3 proteins for polyubiquitination and subsequent degradation through the proteasome (Figure 1.4 and Figure 1.5) (*65, 70*). Structural similarities between Vif and Human Suppressor of Cytokine Signalling 2 (SOCS2), a recognition subunit for the E3 ubiquitin ligase complex, allows Vif to mimic the role of SOCS2 and act as a substrate for receptor for the A3 enzymes (*71, 72*). Upon binding to A3 proteins, Vif acts as scaffold protein for the assembly of a Cullin-5 RING E3 ubiquitin ligase complex. The formation of the complex is initiated by Vif association with an Elongin-B/C heterodimer and the host transcription cofactor CBF-β, followed by the recruitment of Cullin-5 (Figure 1.5) (*73, 74*). The recruitment of the host transcription cofactor CBF-β is an essential component for the stability of Vif and is not directly involved in the Cullin-5 RING E3 ubiquitin ligase complex. However, without CBF-β, the ubiquitin ligase complex cannot form because Vif is thermodynamically unstable and will become degraded in cells unless stabilized by binding to CBF-β (Figure 1.5) (*75-77*). Following complex stabilization, Cullin-5 then interacts with RING finger protein (RBX2), which acts as a docking site for an E2 ubiquitin conjugating enzyme and causes subsequent K48-linked polyubiquitination of A3 proteins (*70, 78, 79*).

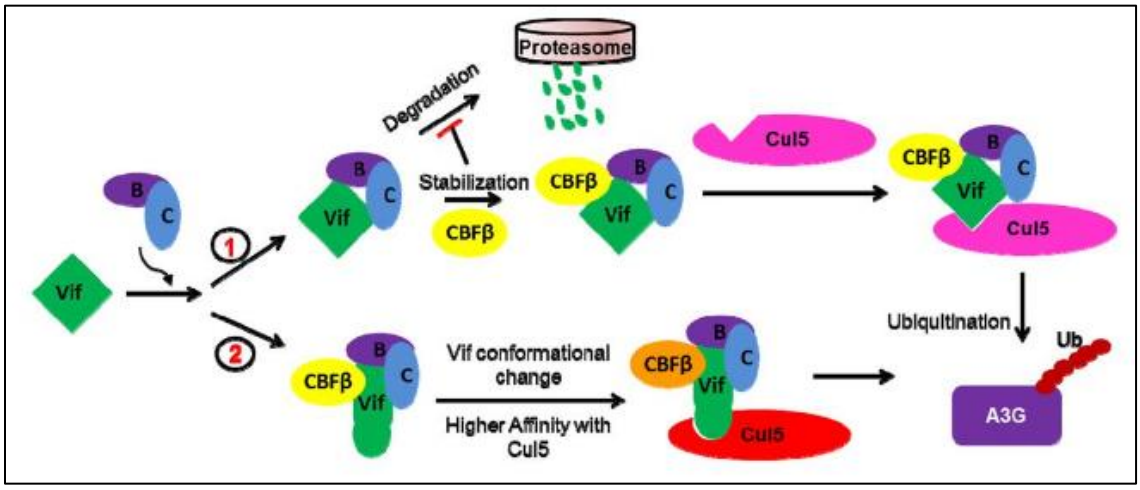


Figure 1. 5 E3 Ubiquitin ligase complex formation by HIV-1 Vif

Proposed mechanism by which Vif mediates proteosomal degradation of APOBEC3 enzymes via a Cullin 5 E3 ubiquitin ligase complex (*78*). Pathway 1 indicates Vif interaction with Elongin B and Elongin C followed by mandatory stabilization with CBF-β. If Vif does not interact with CBF-β, it will become degraded in the proteasome. Pathway 2 indicates successful viral counteraction with Elongin B, Elongin C and CBF-beta leading to conformation change in Vif, which results in a higher affinity to Cullin-5. Both pathways merge after the E3 complex fuses with Cul-5 to form the ligase complex, which can then target A3 proteins for polyubiquitination and result in their degradation through the proteasome pathway. Reprinted with permission from (*78*).

1.7 APOBEC3 Interaction Domains with HIV-1 Vif

The NL4-3 and LAI strains of HIV-1 are examples of once circulating strains from HIV-1 group-M subtype-B that have multiple genomic variations, including areas of their Vif coding regions. Presently both NL4-3 and LAI have been adapted for lab use and through their genetic differences in Vif, they are useful as comparative strains for studying antagonistic HIV-1 responses to A3 enzymes. Namely, LAI Vif has been observed to have a more robust response antagonizing A3H restriction in comparison to NL4-3 Vif. This has been attributed to a single amino acid substitution, at Vif position 48. The substitution of an asparagine (48N) for a histidine (H48) at this position confers resistance to the stable A3H haplotypes (*80*). The Vif activity is exacerbated by alterations in position 39 of the Vif coding region, where the substitution of a Valine (V39) for phenylalanine (39F), confers greater Vif sensitivity to A3H (*80, 81*). Interestingly, N48H and V39F polymorphisms are still prevalent in circulating strains of HIV-1 and appear to be selected for based on prevalence of active/inactive A3H haplotypes within a population (*81, 82*). Active and inactive A3H haplotypes are further discussed in Section 1.8, However, most A3G, A3F and A3D polymorphic variants show no resistance to either V39F and N48H Vif polymorphisms, with even the "weaker" NL4-3 Vif is still able to recognize and degrade A3G, A3F and A3D (*80*). Nevertheless, there are distinct Vif mutations that affect A3G-Vif and A3F-Vif interactions. These sites include Vif positions 14-17 for A3F and Vif positions 40-44 for A3G (*83-86*). Therefore, regardless of the potent A3 antagonism by HIV-1 Vif, the presence of multiple A3 interaction domains and high mutagenic potential of the virus can results in an ongoing coevolution at a population level between A3s and HIV-1 Vif.

1.8 APOBEC3 polymorphisms result in HIV-1 Vif evasion

The HIV-1 Vif is an effective mechanism for defending against deaminase action of APOBEC3 enzymes. However efficient, Vif function does not lead to the complete degradation of the cellular A3s due to observable A3 mutation patterns in HIV-1 proviral DNA through the course of fully competent viral infections (*87*). Mutational footprints of the cytoplasmic A3 enzymes, A3D, A3F, A3G, and A3H, have all been documented in cells infected with partially defective Vif or wild type Vif HIV-1 strains, indicating that this subset of cytoplasmic A3 enzymes can be expressed, encapsidated and catalytically deaminate HIV-1 ssDNA prior to integration even in the presence of Vif (*45, 87, 88*). However, the capacity for A3 mediated retroviral restriction differs between geographical populations, and these modulations have been attributed to numerous single

nucleotide polymorphism (SNPs) present in each A3 gene (*89*). Although all A3 enzymes have characterized SNPs that affect catalytic activity, the most prominent and abundant set of SNPs are in the gene coding for A3H. A3H has seven known haplotypes (hap I-VII) based on five SNPs that result in various amino acid deletions and alternate amino acids (N15Δ, R18L, G105R, K121D, E178D) and an additional four splice variants (SV) of hap I (SV154, SV182, SV183, and SV200) (*81, 90, 91*). The most notable phenotypic changes resulting from A3H SNP variation is the alteration in both protein stability and susceptibility to Vif-mediated degradation (Figure 1.6) (*81, 90, 91*). For example, A3H hap II, hap V, and hap VII are more stable in cells and become packaged in virions to a greater extent in comparison to unstable A3H haplotypes (hap III, hap IV, and hap VI) or partially unstable A3H hap I. This variation of A3H in the population results in variations in HIV-1 Vif ability to bind to A3H, resulting in the ability of some A3H stable haplotypes to restrict HIV-1 in the presence of an active, but "weak" HIV-1 Vif that is similar to HIV-1 NL4-3(*81, 90*). The crucial polymorphisms that results in A3H resistance against Vif is the presence of the R105 SNP in combination with N15 and R18, the trailing SNPs (K121D and E178D) can be variable among the Vif resistant K121D, which occurs in A3H hap VII (*81, 90*). A3H haplotypes are observed to be variable in all populations and are well characterized in African, European Caucasian, and Asian populations; with inactive haplotypes being associated with a greater susceptibility to HIV-1 infection, higher viral loads, greater decline of $CD4^+$ T-cell counts, and more rapid progression to AIDs defining conditions (*81, 90-92*). Although A3H is the only A3 to have characterized resistance to HIV-1 Vif, the presence of circulating viral strains with inactivating Vif mutations and deletions have also been observed, which enables all A3s to affect HIV-1 replication (*93, 94*).

Other notable A3 polymorphisms that effect HIV-1 restriction includes, A3G H186R and A3F A108S/V231I (*95-98*). The A3G H186R polymorphism is variable in African, European Caucasian, and Asian populations, with the 186R genotype being statistically linked to rapid decline of CD4+ T cell counts, increases in viral load and an accelerated progression to AIDS defining conditions (*89, 95*). The allelic frequency of the more robust A3G H186 polymorphism shows a progressive decline in African (0.37), European Caucasian (0.12) and Asian (0.03) populations (*95, 96*). The A3F 231I polymorphism is also associated to similar prognoses to the A3G 186R polymorphism; displaying higher viral loads and a more rapid progression to AIDS defining conditions (*89, 97, 98*). The allelic frequency of the more-active A3F 231V

polymorphism also shows a decline in allelic frequencies in African (0.80), European Caucasian (0.48) and Asian populations (0.30) (*89*, *97*).

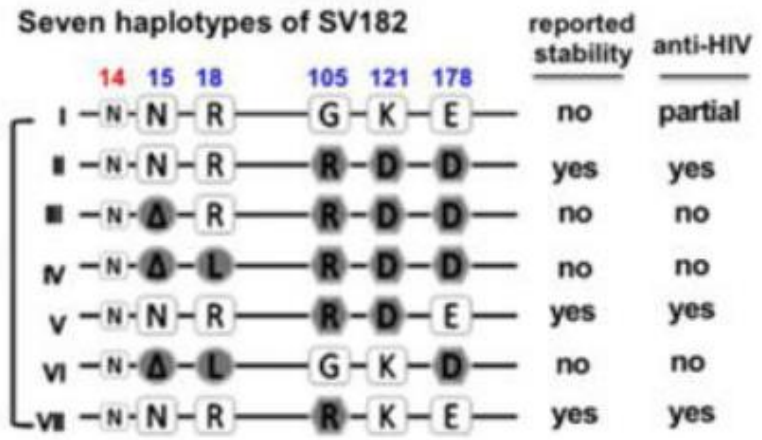


Figure 1. 6 Stability and activity of APOBEC3H haplotypes

Schematic overview of the amino acid changes that comprise the seven known haplotypes of A3H based on the following possible combinations of polymorphisms, N15Δ, R18L, G105R, K121D and E178K. The SV182 is a single splice variant of A3H. The N15 and R105 amino acids are essential for A3H protein stability and only haplotypes II, V and VII contain these amino acids. Reprinted with permission from (*99*).

1.9 APOBEC3 polymorphisms reflect individual human ancestry

Population-based genetic variation of A3 genes have been observed in African, European Caucasian and East Asian lineages, with certain polymorphisms compromising HIV-1 restriction efficiency (*95-98*). These polymorphic variations have also been correlated with human ancestry and appear to dampen when overlaid with classical theories of human ancestral dispersion from central Africa (*100, 101*). That is, according to human fossil records, using single or multiple wave dispersal models, human ancestors were residing and migration within central Africa 300 thousand years ago (Kya), and began expanding into regions of Eastern Europe by 120 Kya, and to the East-Asian seaboard by 70-80 Kya (Figure 1.7) (*100, 101*). Central Africa is also the epidemiological epicenter of HIV-1 evolution, whereby the Simian Immunodeficiency virus from chimpanzees (SIVcpz) transferred zoonotically to humans (*102-104*). Africa has been termed an area of high retroviral challenge due to endemics of not only HIV-1, but also HIV-2, which resulted from a zoonotic transfer of SIV from Sootey Mangabeys (SIVsmm) (*102-104*). The high frequency of retroviral challenge in Africa is thought to have been a primary selective pressure in the development of highly HIV-1 restrictive A3 genes in primates (*6, 15, 89*). In support of this, A3 repertoires of human ancestors that dispersed from Africa to settle in present-day Europe and East Asia exhibit a dampened effect with A3 anti-HIV-1 activity, where time/distance of dispersion from Africa is negatively correlated with HIV-1 restriction efficiency (*89, 95, 96*).

Ancestral dampening of anti-HIV-1 responses is evident in all cytoplasmic A3 enzymes. For example, the more active A3G H186 SNP is most prevalent in African populations (37%), intermediate in European Caucasian populations (13%), and lowest in Asian populations (<1%) (*89, 95, 96*). The more active A3F 231V-108S SNP is highest in African populations (76%), intermediate in European Caucasian's (48%), and lowest in Asian populations (26%) (*89, 97, 98*). Lastly, the N15 and R105 SNP, present in all active A3H haplotypes also has the greatest allelic frequency in African populations (88%), intermediate in European Caucasian's (39%), and lowest in Asian populations (26%) (*81, 90*).

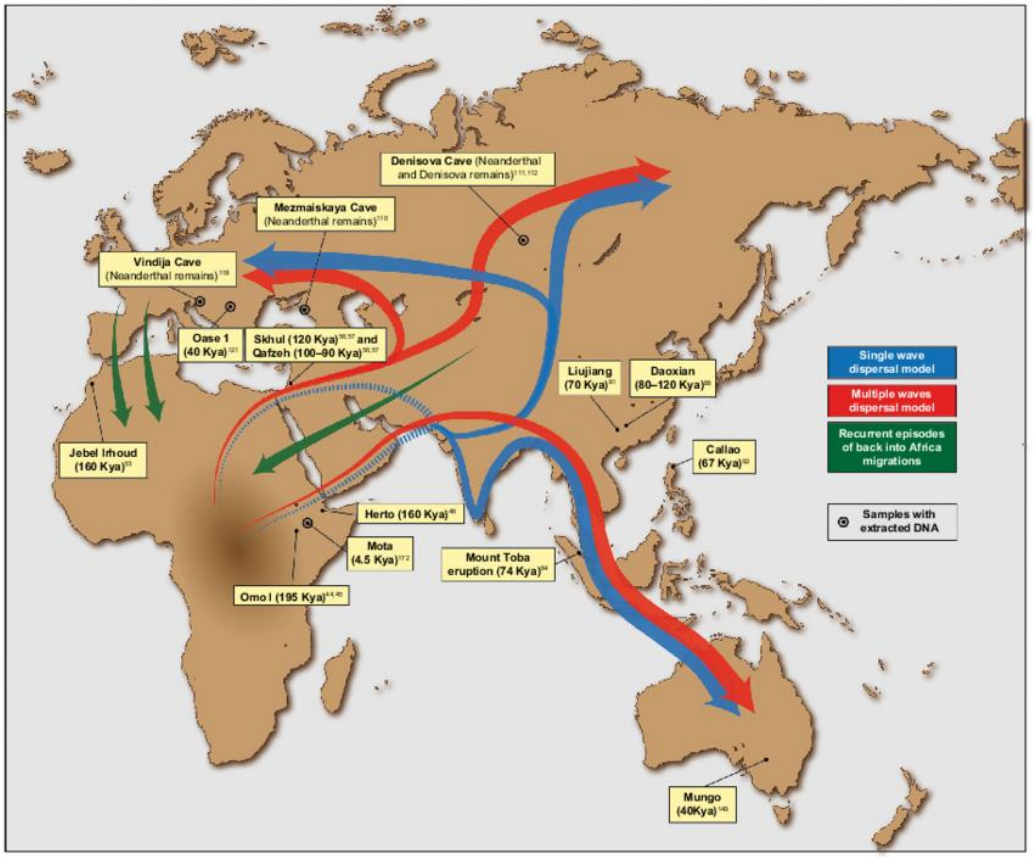


Figure 1. 7 Models for human dispersal from Central Africa

Map of putative migration waves of human ancestors from Central Africa based on fossil recordings at each location. Single wave theory (blue) supports a single dispersal event away from central Africa and into central Europe and the to East Asia or Oceanic states. Multiple wave theory (Red) indicates at least two separate dispersal events separated by 40 Kya. Recurrent theory (green) indicates re-entry states from Western Europe and Central Asia approximately 180 Kya. Adapted from (*100*).

1.10 APOBEC3 Influence on HIV Evolution

The diversity of A3 paralogs observed in human populations provides evidence of a co-evolutionary arms race between A3 viral restriction factors of hominin ancestors and adaptive viral defense measures (*105*). A3 enzymes are just one of many viral restriction factors present in the innate immune system, other constituents include Trim-5 alpha, SAMHD1 and Tetherin (*106*). Each subset of enzymes inhibits a separate stage of the viral lifecycle, for the examples given areas of inhibition include: capsid uncoating (Trim-5 alpha), reverse transcription (SAMHD1), and viral budding (Tetherin) (*107-110*). These restriction factors are all designated as interferon-stimulated genes (ISG), being that viral infections are typically detected by cytoplasmic membrane-bound pattern recognition receptors (PRRs) such as toll-like receptors (TLRs), which trigger an interferon (IFN) response (*106*). Regardless of IFN stimulation, some restriction factors like A3 paralogs are constitutively expressed in both T-cells and germ cells (*22*, *48*). The constitutive expression of A3 enzymes and HIV Vif antagonism has allowed for the development of one of the most pronounced relationships between host and viral defense factors. This interaction helped shape the cross species transfer of simian immunodeficiency virus in chimpanzees (SIVcpz) to HIV-1 in humans (*105*).

HIV-1 resulted from the cross-species transfer to humans from SIVcpz on at least two separate occasions, one of which spawned the group M designated virus responsible for the global AIDS epidemic (*103, 104*). Host restriction factors including A3 enzymes decrease the likelihood of lentiviral transmission from the host, but are usually counteracted on a species-specific manner (*106*). One trade-off alteration in the evolution of SIVcpz Vif allowed for the development of inter-species A3 antagonism at a cost of SAMHD1 susceptibility (*105*). The gene that encodes SAMHD1 antagonism is Vpx, although present in other lentiviruses such as SIVsmm, it is not present in SIVcpz and is theorized to have been deleted during the creation of SIVcpz. The creation of SIVcpz has been identified as being a result of cross-species recombination of two SIV lineages from red-capped mangabeys (SIVrcm) and SIVmus/mon/gsn from Cercopthiecus monkeys (Figure 1.8) (*111, 112*). The recombination event led to the deletion of Vpx and reconstruction of Vif via overlapping its 3' end with the Vpr gene, allowing SIVcpz Vif to antagonize hominid A3s, lowering the barrier for inter-species infection in humans (*105*).

The modification of SIVcpz Vif facilitated the establishment of an infection in human hosts, giving rise to the HIV-1 virus. Due to the lack of DNA repair mechanisms, mutations arise in the HIV-1 genome at a much faster rate than DNA segments of similar size in the human genome this fashion influences evolution at a much faster rate (*113*). The question then arises, how does a host cope in this arms race? The answer comes from the relative size of each respective genome, whereby head-full packaging of lentiviral vectors can only accommodate around 10kb of coding RNA, whereas humans have 3000 Mb of coding DNA (*114*). Having over 30x the coding capacity per strand of HIV allows the human genome to carry duplicates of essential genes such as A3 paralogs and also carry two strand options of heterozygosity and homozygosity. Heterozygosity in restriction factors has been shown to be advantageous to populations of African green monkeys, where monkeys with heterozygous polymorphic variations of A3G disallowed the evolution of SIV Vif to antagonize A3G, while monkeys with both pairs of homozygous alleles resulted in Vif antagonism (*115*). Altogether, current data indicates that maintaining variable polymorphisms in a population is entirely beneficial when countering the offenses of a highly mutagenic virus like HIV-1.

Another mechanism for host competition with viral evolution is gene duplication, with the seven paralogs of human A3s being of prime example. By duplicating genes, not only can the host evolve against multiple different viral targets by increasing copy numbers, but also provide a more diverse defenses against a single target like HIV-1. For example, A3F and A3G paralogs interact with HIV-1 Vif on separate domains, limiting the ability of the virus to evolve (*85, 88, 116*). Similar interactions have also been observed between A3H and HIV-1 on a population level, where A3 polymorphic variations in a population influence Vif polymorphisms (*81*). In populations of higher retroviral challenge such as central Africa, the prevalence of more robust, stable, A3H haplotypes (II, V, VII) increase and where stable A3H haplotypes are more prevalent so is more robust F39 HIV-1 Vif variant (Figure 1.9) (*81*). In areas were unstable A3H haplotypes are prevalent, such as in Asian populations, the less robust V39 HIV-1 Vif variant is more prevalent (Figure 1.9) (*81*).

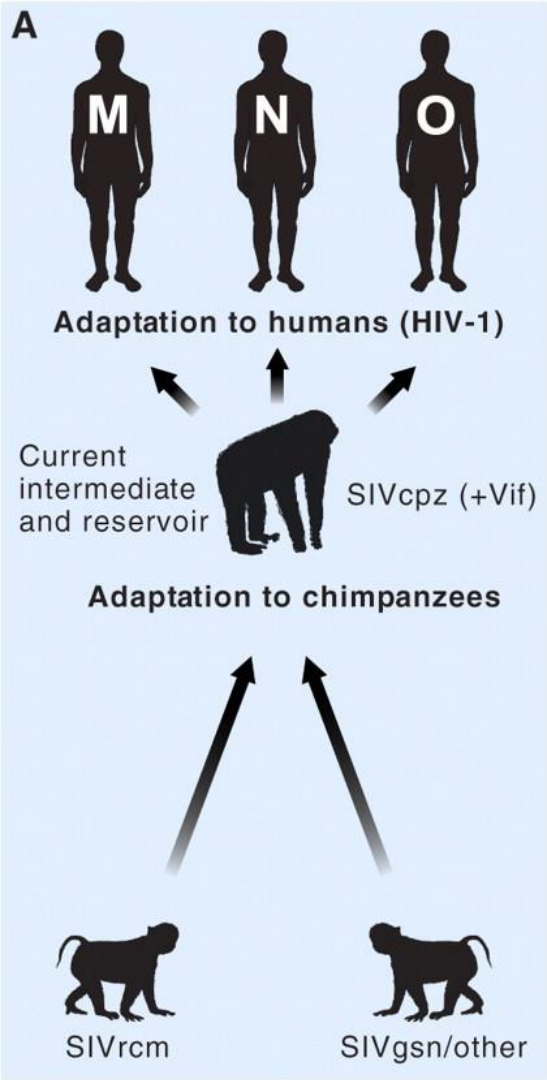


Figure 1. 8 HIV-1 group evolution through recombination of ancestral simian viruses

Diagram depicting the reservoirs of SIVrcm and SIVgsn, which under co-infection and recombination events resulted in the deletion of Vpx and fusion of the 3' end of Vif to Vpu, creating SIVcpz. The modified SIVcpz Vif allows for inter-species A3 antagonism with humans. The lowering of this inter-species barrier to infection ultimately lead to the adaptation of SIVcpz to HIV-1 groups M, N, and O. Reprinted with permission from(*112*).

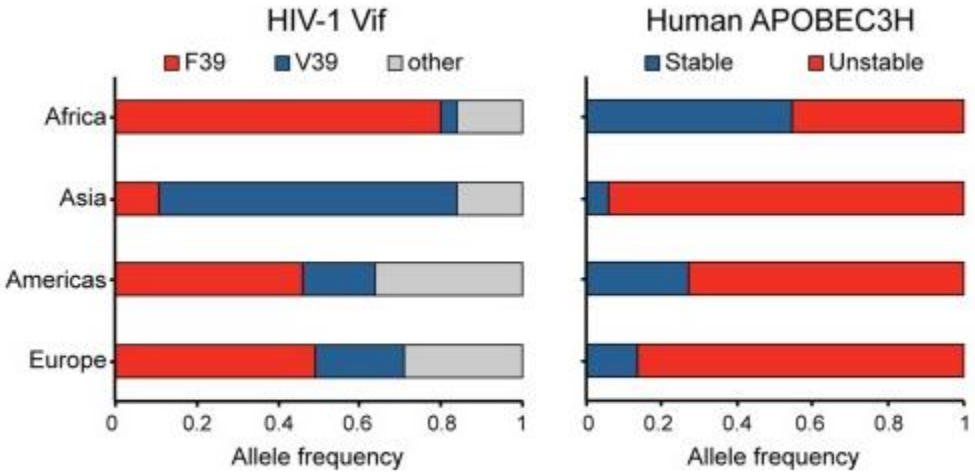


Figure 1. 9 Global distributions of HIV-1 Vif alleles correlated with human A3H haplotypes

The HIV-1 Vif bar chart depicts the frequency of HIV-1 isolates encoding a phenylalanine or valine at Vif residue 39 from the indicated geographic regions (n=9713). The Human APOBEC3H bar chart shows the frequency of stable versus unstable *A3H* alleles from the same geographic regions (n=1092). Reprinted with permission from (*81*).

1.11 Effects of APOBEC3 co-expression in cells

Within the A3 family there is a cascade of deaminase activity, with A3G being the most potent restrictor of ΔVif HIV-1, followed by A3F, A3H, and A3D respectively (*45, 46, 49, 53, 70, 83, 117, 118*). However, from quantitative analysis in human PBMCs and CD4+ T-cell lineages we know that all cytoplasmic A3s are constitutively expressed and can be upregulated when stimulated with interleukins and interferons associated with cell stress and viral infection (*22, 48, 119, 120*). A3s also share a great extent of nucleotide consensus and have been shown to self-oligomerize into higher order complexes (*33, 51, 52, 55, 56*). Additionally, A3G and A3F have been shown to hetero-oligomerize when co-expressed and have a cooperative effect on HIV-1 proviral hypermutation and restriction (*55*). When A3F and A3G are co-expressed in cells, there is a greater decline in ΔVif HIV-1 infectivity and two-fold greater levels of A3F and A3G -specific mutations in comparison to A3F or A3G expressed alone (Figure 1.10) (*55*). The increased mutational frequency observed in the A3F and A3G dual expression condition also results in a higher frequency of virus inactivating mutations indicating an inter-A3 cooperative that compounds viral restriction (Figure 1.10). The study used size exclusion chromatography, rotational anisotropy, and co-immunoprecipitation to show that A3F and A3G directly interact, without a nucleic acid intermediate, and form a hetero-oligomer that and can more efficiently scan for deamination sites on ssDNA with intervening RNA-DNA hybrids using both sliding and jumping ssDNA scanning mechanisms (*55*).

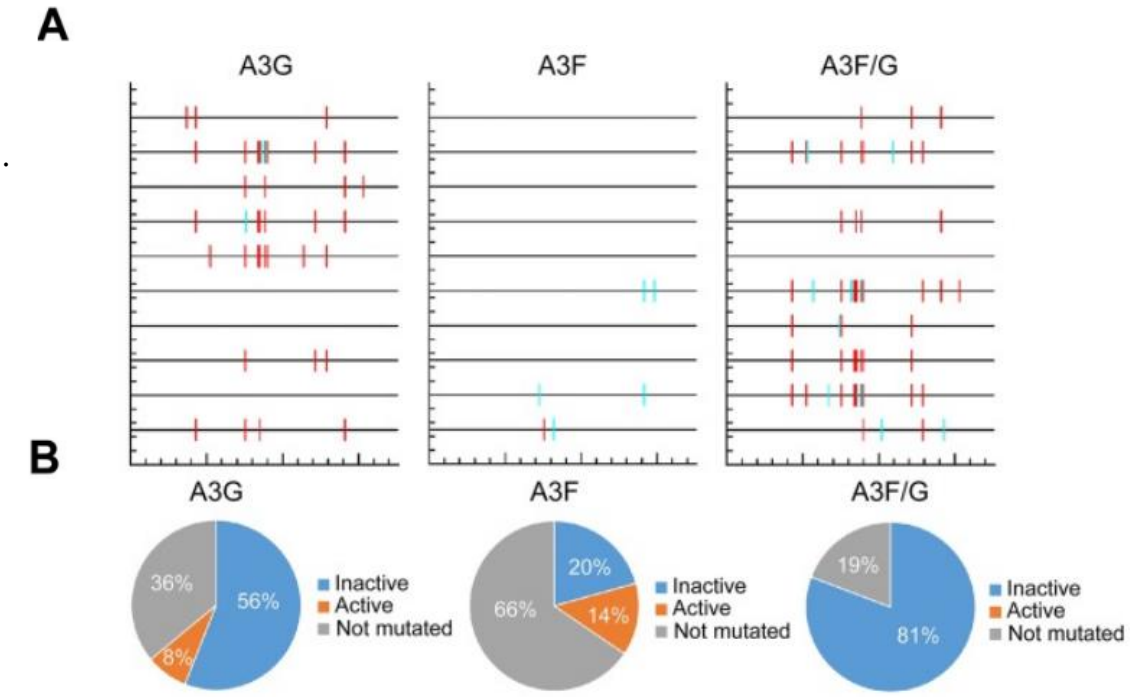


Figure 1. 10 Effects of APOBEC3G and APOBEC3F co-expression in cells

Co-expression of A3F and A3G is shown to result in co-encapsidation and co-mutation of ΔVif HIV-1 proviral DNA. (A) Mutation spectrum plots using HIV-1 protease sequences isolated from 10 individual clones generated with the expression of A3G, A3F or A3G/A3F expression vectors. A3G specific mutations (GG→GA) are shown in red and A3F specific mutations (GA→AA) are in cyan. (B) For each A3F, A3G and A3F/A3G, individual amino acid changes were analyzed for mutations resulting in an active HIV-1 protease (orange), inactive HIV-1 protease (blue) or a non-mutated protease (grey). Reprinted with permission from (*33*).

2.0 OBJECTIVES AND HYPOTHESIS

2.1 Rationale for Hypothesis

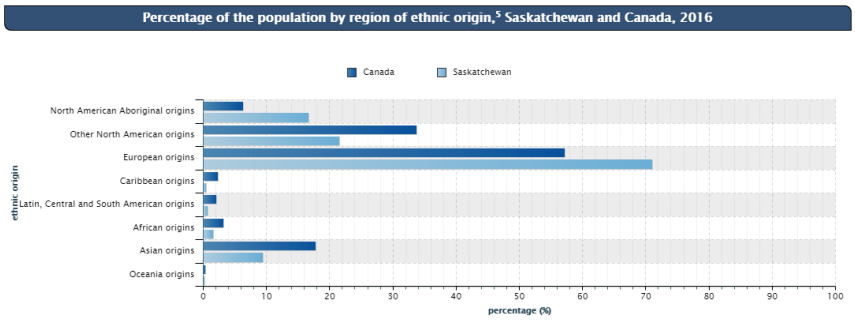
Current literature supports cause for ancestral ties to A3 repertoire variations in global populations, with activity declines correlating with human ancestral dispersal from Central-Africa. This trend has been attributed to a decline in positive selective pressures maintaining robust retroviral restriction factors; with pressures being frequency of retroviral challenge from endemic SIV, HIV-2 and HIV-1 present in Central Africa. The effects of physical human dispersal away from an area of high retroviral challenge thereby decreases the selective pressure to maintain robust A3 repertoires. (*41, 100-102, 121, 122*).

Current studies examining A3 genotype variation between populations focus on single ethnic lineages, individual expression of A3s, and analysis bases on allelic prevalence of A3 SNPs. Being current with the times and acknowledging the current state of globalization indicates that most populations have diversified from historic migration patterns of our human ancestors and that ethnic lineages reflect mix populations. We currently do not know the effects of A3 polymorphic variations on a mixed population and how variation in A3s affect inter-A3 interactions when expressed in combination.

HIV-1 affects more than 85,000 Canadians with mortality rates of 1,750 individuals per year since the advent of blood screening tests developed in 1983. There were 2,500 new infections in Canada in 2017, a steady increasing trend and an estimated increase to 3,700 infections by 2018 (*123*) , HIV-1 is a growing public health concern for *all* Canadians and novel methods for treatment and resolve are of the greatest priority. The province of Saskatchewan is an extraordinary example, having incidence of HIV-1 infections over 3-fold greater than the national average of 5.4/100,000 population, providing the province with the highest HIV-1 prevalence in Canada. Even more concerning is the infection prevalence in rural populations of the province; where disease prevalence distends to over 14x the national average. These values are comparable only to endemic areas in Central Africa. The rise of HIV-1 prevalence in the province is attributable to socioeconomical, demographic and genetic factors. However, due to current states in globalization, population diversity may be altering innate mechanisms for restricting retroviral challenge. The current demographics of Saskatchewan indicate a mixed-population consisting of European Caucasian (70.9%), Indigenous (16.6%) and East Asian (9.3%) (*124*). We expect the SNP

repertoires will reflect this demographic, resulting in a population genetically more susceptible to HIV-1 infection than African populations.

A



B

Total population by Aboriginal identity and Registered or Treaty Indian status, Saskatchewan, 2016 Census

Aboriginal identity	Number	Percent (%)
Total – Population by Aboriginal identity	1,070,560	100.0
Aboriginal identity	175,015	16.3
Single Aboriginal response	172,810	16.1
First Nations (North American Indian) single identity	114,570	10.7
First Nations single identity (Registered or Treaty Indian)	106,440	9.9
First Nations single identity (not a Registered or Treaty Indian)	8,130	0.8
Métis single identity	57,880	5.4
Inuk (Inuit) single identity	360	0.0
Multiple Aboriginal identities	1,300	0.1
Aboriginal identities not included elsewhere	905	0.1
Non-Aboriginal identity	895,540	83.7

Figure 2. 1 Statistics Canada 2016 census report

Population census summary from 2016 displaying percentage of population of Canada, overlaid with Saskatchewan percentage of population. (B) Population census summary from 2016 displaying population values of Saskatchewan, percent Indigenous population and percent Indigenous identity.

2.2 Experimental Questions

1. How variable are APOBEC3 (A3) genotypes in Saskatchewan?
2. Do variable A3 genotypes affect HIV-1 restriction when expressed in combination?

2.3 Hypothesis

- We hypothesize, that due to global diversification and variations in regional selective pressures for A3 SNP maintenance, A3 repertoires will be *unique* in Saskatchewan populations but will reflect current demographics (European Caucasian and East Asian).

2.4 Project objectives

The experimental strategy can be broken up into five main objectives:

1.) Genotype A3 repertoires from donor populations in Saskatchewan to determine prevalent A3 subsets in the mixed population of Saskatchewan.

2.) Use population-based genotypes as a template for A3 expression vector design for use in single-cell infectivity assays to determine if A3 interactions influence HIV-1 restriction in the presence and absence of HIV-1 Vif.

3.) Examine cellular expression and viral encapsidation frequencies of genotype derived A3 repertoires when co-expressed.

4.) Sequence circulating HIV-1 strains from HIV-1+ donors in the province of Saskatchewan.

3.0 Genotyping of Patient Sample PBMCs

3.1 Introduction

Genotyping our PBMC donor set was an essential component in determining the A3 repertoire in the Saskatchewan population. A significant amount of time was spent in the development and troubleshooting of these foundations. Troubleshooting was not limited to the restriction digest genotyping assay, but it was certainly a large contributor, and ultimately due to inconsistencies in data outputs this assay was replaced with more accurate Sanger sequencing genotyping. The end-product being the genotyping of the complete A3H loci for 13 individual donors in our population, consisting of the compilation of five SNPs, producing one of seven known haplotypes.

3.2 Materials and Methods

3.2.1 Donor PBMC and DNA isolation

PBMCs were isolated from whole blood samples using Ficoll-Paque (GE-Healthcare) density gradient centrifugation in Sepmate50 tubes (STEMCELL Technologies). DNA was isolated from donor PBMCs using a PAXgene Blood DNA Kit (Qiagen).

3.2.2 Restriction Digest Genotyping

Methods for A3H genotyping were drawn from Wang *et al. (90)*, and adapted for A3F, A3G, and A3D genotyping. Previously identified polymorphisms resulting in A3H (R18L, G105R, K121D, and E178D, E140K), A3F (A108S and 231IV), A3G (H186R) and A3D (R97C) amino acid changes were genotyped using restriction enzyme digestion of amplified genomic DNA from human donors (*90*). Each experiment and selected restriction enzymes were designed around naturally occurring restriction sites that can distinguish and cut at polymorphic alleles. Digestion positive controls were designed so that the same enzyme that cuts the polymorphic site also cuts a non-polymorphic or conserved region on both alleles creating a secondary band digestion, this verifies complete digestion. Restriction enzymes, primer sequences, PCR condition, and banding patterns are shown in Appendix Tables 1, 2 (A-I), and 3. Due to a lack of control restriction sites for A3F (A108S and 231IV), A3G (H186R) and A3D (R97C) SNPs within the amplicon, separate double-stranded (ds) DNA-based controls were designed to ensure complete restriction enzyme digestion. Each control was a 30 nt or 40 nt synthesized double-stranded oligonucleotide with each restriction site used for genotyping. Enzyme digestion with each control was done in tandem with

donor digestions. Primer sequences and restriction enzymes are shown in Appendix Table 4. The only polymorphism that could not be determined via restriction digestion was the A3H N15Δ deletion. The deletion was detected by resolving the PCR product on a 15% urea-PAGE gel and the amplicon size determined the presence or absence of a deletion.

3.2.3 Sanger Sequencing genotyping

The A3G, A3F, A3D, and A3H polymorphic sites were amplified using PCR. The A3G H186R polymorphism was detected using the following primers: (forward) 5′ AATTTCAGCACTGTTGGAGC and (reverse) 5′ GAGAATCTCCCCCAGCAT. The A3F V231I polymorphism was detected using the following primers: (forward) 5′ AGCCTATGGTCGGAACGAAA and (reverse) 5′ CTGGTTTCGGAAGACGCC. The A3D R97C polymorphism was detected using: (forward) 5' GTGTATTTCCGGTTTGAGAACC and (reverse) 5' GATGGTCAGGGTGACATTGG. The five A3H polymorphism detection primers are listed in Appendix Table 1. The PCR amplification used Q5 High Fidelity polymerase (New England Biolabs) and cycling conditions were 95° C for 5 min followed by 45 cycles of 95° C for 10 sec, 52° C for 20 sec, 72 ° C for 20 sec. Amplicons were cloned into pJET1.2 using the CloneJET PCR cloning kit (ThermoScientific). Amplicons were Sanger sequenced at the National Research Council of Canada (Saskatoon, Saskatchewan). The resulting A3 sequences were translated using the Expasy Translate tool to identify polymorphisms (*125*)(*125*)(*125*)(*125*) and clones from each donor were analyzed for each A3 condition and their respective polymorphisms to determine homozygosity or heterozygosity of each allele in combination with allelic frequency.

3.2.4 Proviral Sequencing

For proviral sequencing, 1 x 10^5 293T cells were infected with 40 μL of supernatant containing virus in the presence of 8 μg/mL polybrene. The plates were spinoculated at 800 x *g* for 1 h. Cells were harvested after 48 h by removing the media, washing with PBS, and lysing the cells and extracting DNA with DNAzol (Invitrogen). The PCR amplification of the protease region of HIV-1 (351 bp) was conducted as previously described Ara *et al.* 2014 (*50*). Sequences were analyzed with Clustal Omega and Hypermut software programs (*126*) (*127*).

3.2.5 Typing integrated provirus from HIV-1^+ Donors

Using the Qiagen Purgene Cell Kit A, DNA was extracted from 2-4 x 10^6 PBMCs. This kit was utilized because some of the PCRs required large reads of up to 6.5Kb and the kit is able to

minimize fragmentation up to 200Kb. Prepped DNA samples we subjected to nested PCR protocols using Platinum Taq High Fidelity polymerase (Life Technologies) as described in Bruner et al. (*87*). The outer PCR product amplifies nucleotides 623 to 9686 of the HIV-1 HXB2 strain genome. Then, 1uL of the outer amplification is used in each nested reaction. Detailed PCR protocols and primer sets are Appendix Table 5 and Appendix Figure 1 and were adapted from Buner et al. (*87*). Each product was resolved using agarose gel electrophoresis. The PCR products were then purified using the QIA quick Gel Extraction kit followed by cloning and Sanger Sequencing.

3.2.6 Typing circulating virus

Protocol and primers were from Brumme *et al.* 2011 and Grossman *et al.* 2015 (*128, 129*). Virus is isolated from cell free patient plasma using 0.25-micron filter syringes, followed by RNA extraction using QIAamp viral RNA extraction kit. RNA quantity and purity were determined using a NanoDrop (Thermo Scientific, DE, USA) and samples were stored at -80 until required. First strand cDNA synthesis was performed using random hexamer PCR primers with 1 µLSuperScript III Reverse transcriptase (Invitrogen), 1 µL of 10mM dNTPs 13 µLPCR grade/RNase Free water (Roche), 1 µL DTT and 5 µL of RNA template. cDNA synthesis is facilitated by heating at 65^{o}C for 5 min followed a 2 min incubation on ice. A second master mix containing 4 uL of 5x First-Strand Buffer (375 mM KCl, 250 mM Tris-HCl, 15 mM $MgCl_2$), 1 µL 0.1 M DTT, 1 µL of 40U/mL RNase inhibitor (Roche) and 2 mL of SuperScript-III RT (200 U/mL). The final 20 µL reaction mix is then incubated at 25^{o}C for 5 min followed by 55^{o}C for 1 h and finally 70^{o}C for 10 min to terminate the reaction. Amplification of the target region of RT is facilitated by nested touchdown PCR reactions using primers sets detailed in Appendix tables 6 and 7.

Ethics Statement. This study has been reviewed and approved by the Research and Ethics Board at the University of Saskatchewan (Bio# 14-62 and 13-293). Peripheral Blood Mononuclear Cells (PBMCs) were isolated from de-identified individuals Donors were either HIV+ and recruited through the Positive Living Program (PLP) at the Royal University Hospital or West Side Clinic in Saskatoon or HIV- Donors (HD).

3.3 RESULTS

3.3.1 Isolation of HIV-1 proviral DNA from donors

Through collaboration with Dr. Steve Sanche and Dr. Kris Stewart we began receiving blood donations from consenting patients. We received nine HIV-1^+ blood samples in total, termed PLP-001 to PLP-009, which we used for both A3 repertoire genotyping and proviral typing. Here we used a protocol and primer set from Bruner *et al.* 2016 (*87*) to amplify near-full-length HIV-1 provirus from DNA isolated from donor PBMCs. After multiple rounds of template dilutions, we were unable to amplify any of the six nested fragments using the published protocol (Figure 3.3.1A-B).

3.3.2 Isolation of circulating virus from HIV-1^+ donors

Due to setbacks in HIV-1 proviral isolation from donor PBMCs we then attempted to isolate viral RNA directly from donor plasma to be amplified using RT-PCR. We chose this methodology as a secondary outlet for typing circulating virus because most of the donors were on anti-retroviral therapies (ART) and had undetectable viral loads. The only exception was PLP-003, which had a viral load of approximately 20,000 copies/mL. From the literature, we found that copy numbers above 3000/mL would most likely have yields of enough RNA for cDNA synthesis and subsequent PCR reactions (*129*). Here we used primer sets and protocols from Brumme *et al.* 2011 and Grossman *et al.* 2015, to isolate amplify HIV-1 RNA using RT-PCR (*128, 129*). Because we were only able to obtain 10 mL of blood from our HIV-1^+ donors, the filtered plasma resulted in 2 mL of virus containing supernatant, in which we expected between 20,000 and 40,000 HIV-1 copies. However, the RNA prep resulted in low yields (21 ng/µL). RNA was added to the cDNA synthesis reaction in a total amount of 200 ng. The primary nested PCR primers used for HIV-1 RT amplification from cDNA were 5'CP1 (F) + RT3.1 (R), or PRTO-5 (F) +Out3 (R), the secondaries being 2.5 (F) + RT3798R (R) or 2.5 (F) + 2.3 (R) (Appendix Table 6). As expected, the primary nested amplification resulted in a full band-length smear when ran on a 2% agarose gel (Figure 3.3.1A), however, even with extensive modifications to PCR protocols and template concentrations, we were unable to optimize the assay to amplify HIV-1 RT from circulating HIV-1 from donor PLP-003 (Figure 3.31B-D). The only exception was the sample in Figure 3.31C (lane 2) that contained 10 µL of the outer amplification template and used primer set one (5'CP1 (F) + RT3.1 (R) and 2.5 (F) + RT3798R (R)), Appendix Table 6). This band was gel purified and

selected for cloning, but cloning reactions failed to yield colonies for sequencing. We started with this aim of the project with the intent to isolate a virus circulating in Saskatchewan and clone the sequence into a vector so that it could be used for experiments. However, since we were unable to amplify an HIV-1 sequence, for all further experiments, lab strain virus HIV-1 LAI was used.

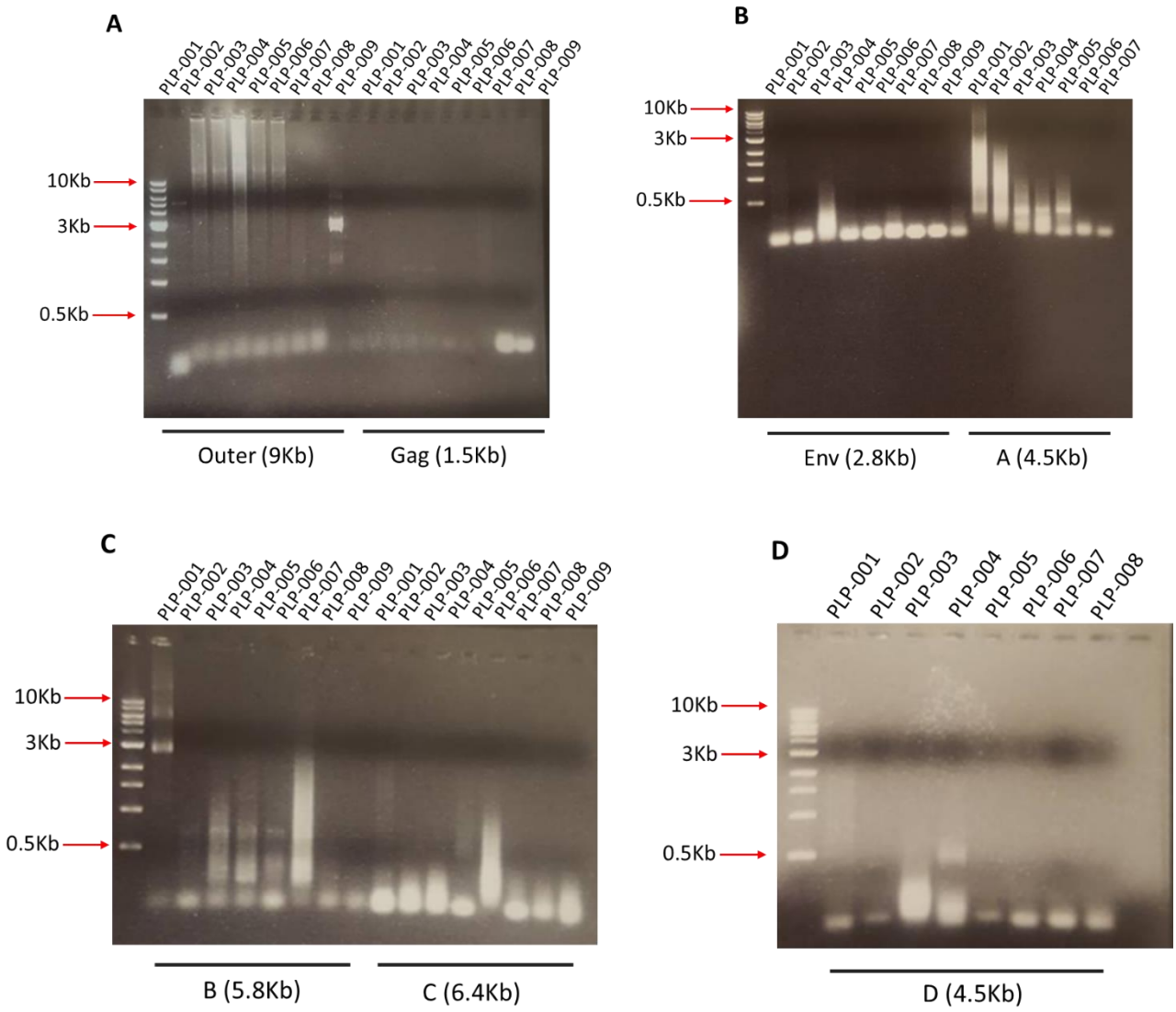


Figure 3.3. 1 Near-Full-Length HIV-1 proviral amplification from donor PBMCs

(A) Lanes 2-10 show the outer 9Kb PCR amplifications for donor samples PLP-001 to PLP-009. The outer amplification was used as a template for all nested PCR reactions. Lanes 11-20 show the 1.5Kb amplification of HIV-1 gag for PLP-001 to PLP-009. (B) Lanes 2-10 show the 1.8Kb amplification of HIV-1 env for PLP-001 to PLP-009. Lanes 11-17 show the 4.5Kb amplification of HIV-1 'A' segment for PLP-001 to PLP-007. (C) Lanes 2-10 show the 5.8Kb amplification of HIV-1 'B' segment for PLP-001 to PLP-009. Lanes 11-20 show the 6.4Kb amplification of HIV-1 'C' segment for PLP-001 to PLP-009. (D) Lanes 2-9 show the 4.5Kb amplification of HIV-1 'D' segment for PLP-001 to PLP-008.

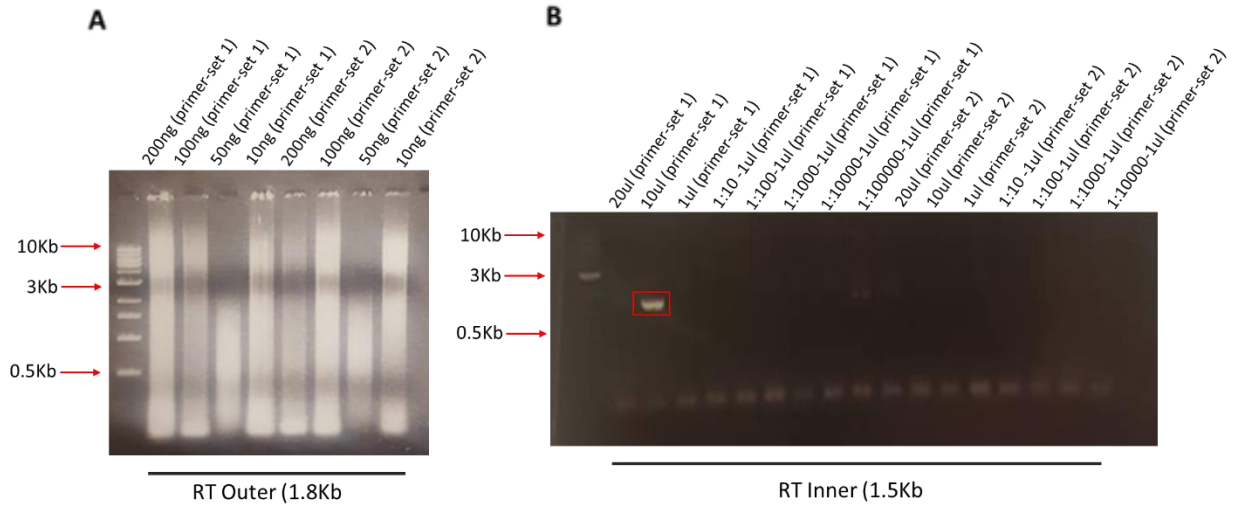


Figure 3.3. 2 HIV-1 RT amplification from circulating virus from donor PLP-003

(A) Outer amplification using primer set 1 (5'CP1 (F) + RT3.1 (R)) and primer set 2 (PRTO-5 (F) +Out3 (R)). (B) Nested RT amplifications using titrated templates from outer PCR amplifications. Template titrations range from 20 µL down to 1:10,0000. Nested primers included: 2.5 (F) + RT3798R (R) and 2.5 (F) + 2.3 (R).

3.3.3 Restriction Digest Genotyping

The same DNA from donor PBMCs used for our attempts at proviral amplification and viral typing was also used for A3 repertoire genotyping. In anticipation of receiving high volumes of donor samples we designed the A3 genotyping to be through restriction digestion of amplified polymorphic sites from donor DNA, a process that was more cost effective than Sanger Sequencing. The A3H genotyping protocol was used from Wang *et al.* 2010 (*90*). The A3G, A3F, A3D genotyping assays were designed for this project. A3 amplification primers, banding patterns, and restriction enzymes are detailed in Appendix Tables 1, 2, and 3, respectively. If multiple SNPs were to be analyzed on the same exon and were in close proximity (A3F 231IV/A108S) single primer pairs could be used for multiple SNP typing. Alternatively, if multiple SNPs were to be typed on different exons, separate primer pairs were required, such as for A3H.

One component of the genotyping assay that was initially overlooked during A3G, A3F and A3D assay design was the presence of an internal control. That is, upon running the first set of gels for A3G genotyping, all present bands ran at approximately 113 bp, which we designed as the 'uncut' condition indicating homozygosity for H186, whereas, homozygosity for R186 would result in the 'cut' condition resulting in 90 bp and 23 bp segments. As a result, it was impossible to differentiate between the 186H homozygosity and a faulty restriction digest. We confirmed a sample being uncut by running the undigested PCR product next to the donor digests (Figure 3.3.2A). The band patterns were identical, indicating the requirement for restriction digest design to cut at both the polymorphic site and in a non-polymorphic region to ensure complete digestion.

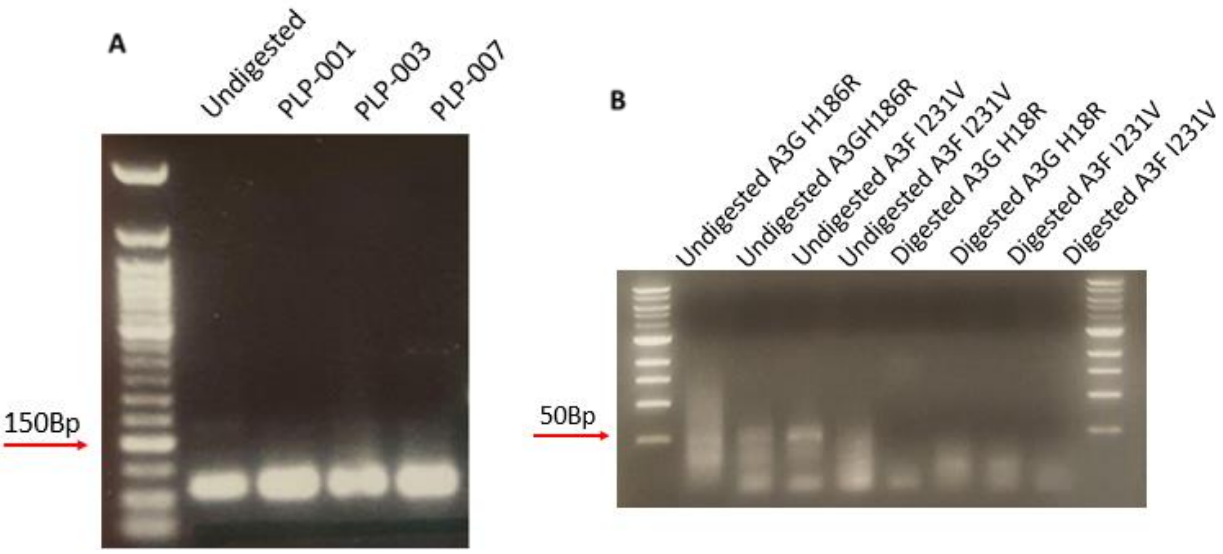


Figure 3.3. 3 APOBEC3G and APOBEC3F restriction digest genotyping assays and restriction controls

A) A3G exon amplification and FSPI digestions for H186R genotyping. Three donor samples were digested (PLP-001, 003 and 007) and ran on a 2% agarose gel and stained with ethidium bromide, along with an undigested control. (B) Synthesized oligonucleotides (40 nt) used for A3G, A3F and A3D control digests in restriction digest genotyping assay. A3G H186R and A3F 231IV were used as examples following FSPI and BSmaI digestion, respectively.

Unfortunately, restriction mapping for dual polymorphic and non-polymorphic control site cutters for A3G, A3F and A3D proved futile. That is, we could not identify restriction enzymes that would cut with single nucleotide specificity at each polymorphic site and at another upstream or downstream conserved site for a control. As a final attempt to incorporate an internal control into the A3G, A3F and A3D digests, a synthetic oligonucleotide-based control method was employed. 30 nt or 40 nt dsDNA oligonucleotides were designed using the NEB cutter program (*130*) and contained the respective restriction sites for each A3 SNP digest. The 30 nt or 40 nt complementary oligonucleotides were annealed using a high salt annealing buffer at 95 ^{o}C for 5 min, then cycled down 1 ^{o}C every min to 25 ^{o}C and then digested using the same enzymes used for their respect A3 paralog, as detailed in Appendix Table 3. Enzyme activity was determined by the production of a 20 nt product in a control digest. After doing multiple sample control digests, the resolution of each banding pattern could not be increased to produce a reliable control. Input DNA concentrations, restriction enzyme concentrations, loading volumes and gel densities were all modulated. The product with the best resolution still left reasonable error (Figure 3.3.2B). Digested oligonucleotides did appear to shift lower than the undigested controls, but were still not discernible between 40 nt and 20 nt.

Challenges with poor resolution, sample laddering and uncertain controls were a combined cause to stop the restriction digest genotyping assay and confirm and extend results with Sanger sequencing-based genotyping. However, due to the amount of time spent troubleshooting the restriction digest assay we decided to streamline the number of SNPs analyzed using the Sanger method. We still chose to examine all five A3H SNPs, A3G H186R and A3F 231IV SNPs, but A3D R97C was abandoned due the low capacity of A3D to restrict HIV-1, even with the more-active Arginine (R) at position 97 (*89*).

3.3.4 Sanger sequencing-based genotyping of A3H

Using the same A3H primers used in the restriction digest genotyping assay, A3H was re-amplified from donor PBMC DNA, ligated into a pJet cloning vector, transformed into DH5-alpha E. *coli* cells and plated on LB-Amp plates for individual clone selection. Plasmid DNA from each clone set were then prepared for sequencing using supplied CloneJET primers. For accurate determination of the complete A3H genotype, 6-8 clones per A3 SNP were sequenced to ensure

probable detection of both alleles and determination of homozygosity and heterozygosity. When the A3H allelic frequency was analyzed a great diversity even in the small population set of 13 donors was observed (Figure 3.3.3A-B). Specifically, we observed allelic frequencies of hap I (35 %), hap II (4 %), hap III (25 %), hap IV (15 %), hap V (8 %), hap VI (8 %) and hap VII (5 %). Based on previously published data on A3H haplotype activity, A3H activity in our population was at 30% (Figure 3.3.3C), which is similar to combined data sets of larger population-based studies (*90, 91*). That is, in Wang *et al.* 2011, which looked at isolated individual populations showed active A3H haplotypes occurring at 44.4 % in European Caucasian lineages and at 25.0 % in Chinese lineages, averaging at 26.0 % (*90*).

A

Patient Sample	Allele 1	Allele 2	Haplotypes		Activity
HP-002	NRGKE	ΔLRDD	HapI	HapIV	inactive
HP-003	ΔRRDD	NRGKE	Hap III	Hap I	inactive
HP-004	ΔRRDD	ΔLRDD	Hap III	Hap IV	inactive
HP-005	NRRDE	ΔRRDD	Hap V	Hap III	active
HP-006	NRRDE	NRGKE	HapV	Hap I	active
PLP-001	NRGKE	ΔLGKD	Hap I	hap VI	inactive
PLP-002	NRRDD	ΔRRDD	Hap II	Hap III	active
PLP-003	NRGKE	ΔLRDD	Hap I	hap IV	inactive
PLP-004	NRRKE	ΔLRDD	Hap VII	hap IV	active
PLP-006	NRGKE	ΔLGKD	Hap I	Hap VI	inactive
PLP-007	NRGKE	ΔRRDD	HapI	HapIII	inactive
PLP-008	ΔRRDD	ΔRRDD	hap III	hap III	inactive
PLP-009	NRGKE	NRGKE	Hap I	HapI	inactive

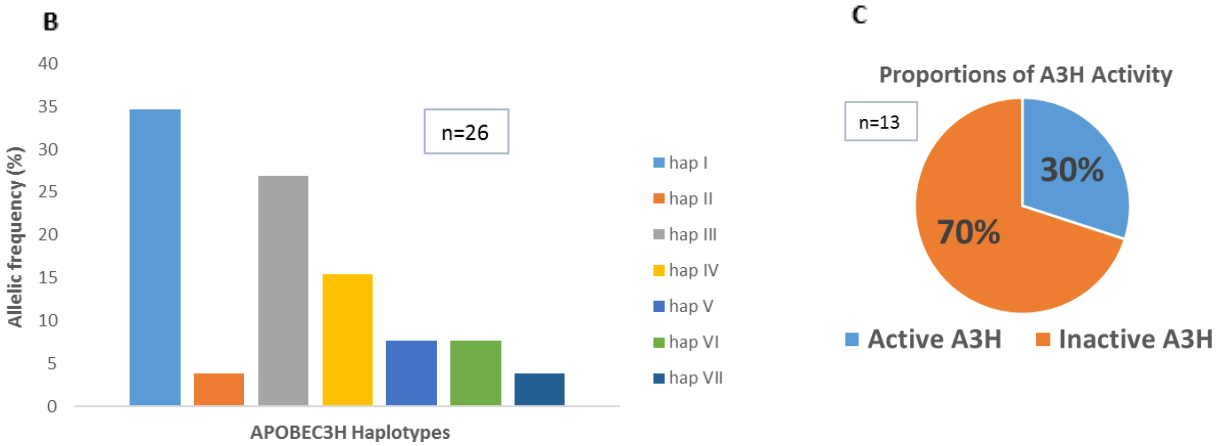


Figure 3.3. 4 APOBEC3H genotype analysis

(A) Allelic frequencies, whole genotypes and summarized haplotypes from 13 donor PBMC DNA samples (B) Summary of haplotype variation as a function of allelic frequency from 13 donor PBMC DNA samples. (C) Proportions of active A3H haplotypes (hap II, hap V, and hap VII) in comparison to inactive A3H haplotypes (hap I, hap III, hap IV and hap VI) from 13 donor PBMC DNA samples. Activity of A3H haplotypes was previously published (*90, 91*).

3.4 Discussion

3.4.1 Development, alteration, and completion of APOBEC3H Genotyping Assays

In Chapter 3, the foundations for my project were laid through a series of trials and tribulations. The first component of the project was to process blood samples for both viral sequencing and A3 and genotyping. Viral typing was an important feature in designing a more physiologically relevant infection model when using Saskatchewan-based donor sets. Our end goal was to characterize circulating strains, identify RT-based drug resistance and Vif mutation variations, and clone infectious molecular clones into plasmid expression vectors to be used in expression and infectivity assays. However, after months of troubleshooting HIV-1^{+} donor proviral amplification and circulating virus isolation with little success we chose to sacrifice this component of the project and use the field standard LAI strain HIV-1 for expression and infectivity assays.

More issues plagued the start-up of donor A3 genotyping when using an adapted assay from Wang *et al* 2011, which did not enable accurate and interpretation of A3H genotypes from restriction digests (*90*). Sanger-sequencing proved to be a reliable and attainable source for the limited donor samples we received. From the data in Section 3.3, we concluded that A3H genotypes are incredibly diverse in our mixed-population set and the combined genotypes resulted in a 30% frequency of active A3H haplotypes. When compared to data presented in Wang *et al* 2011 and Duggal *et al* 2013, we observed that the frequency of A3H activity in our population correlated to a mixed population of European Caucasian and East Asian, which are 70.9% and 9.3% respectively, and is representative of current demographics in Saskatchewan (Figure 2.1) (*89, 90, 124*). On a population level, low frequencies of active A3H, similar to that of our present population have been linked to a greater susceptibility to HIV-1 infection, lower CD4+ T-cell counts through the course of infection, higher viral loads and more rapid progression to AIDS (*92*).

4.0 Preface to studying polymorphic variants of A3F

The data collected in Chapter 3 provided a foothold for examining other A3 paralogs in our population. As seen in Figure 3.3.4A-C a great diversity in A3H genotypes was observed with relative activity presenting at 30%, which correlated to current population demographics in our population of mostly European Caucasian and East-Asian, which have A3H activity profiles of 50% and 11% respectively. With current population percentages taken into consideration, the predicted A3H activity profile correlates to the experimental values. Because we saw little variation between predicted and expected values for A3H activity in our population, we decided to shift our attention to A3G and A3F SNP variations. We wanted to investigate A3G and A3F SNPs together due to previous evidence supporting their co-expression, interaction and increases in HIV restriction when hetero-oligomerized (33). Further, we wanted to determine if common A3G and A3F SNPs affect these interactions and what the results would be in relation to HIV-1 restriction capacity.

5.0 Polymorphic variants of the cytidine deaminase APOBEC3F cooperate to restrict HIV-1

Tyson B. Follack[a], Nazanin Mohammadzadeh[a], Kris Stewart[b,c,d], Stephen Sanche[b, d], and Linda Chelico[a, d]

[a]University of Saskatchewan, Biochemistry, Microbiology, and Immunology, College of Medicine, Saskatoon, Saskatchewan, Canada

[b] University of Saskatchewan, Department of Medicine, College of Medicine, Saskatoon, Saskatchewan Canada

[c]Saskatchewan Infectious Disease Care Network, Saskatoon, Saskatchewan, Canada

[d] Saskatchewan HIV-1/AIDS Research Endeavour, Saskatoon, Saskatchewan, Canada

Experiments that resulted in Figure 5.4, Table 5.2, Figure 5.5, and Table 5.3 were performed by Tyson Follack. Experiments that resulted in Figure 5.1, Table 5.1, Figure 5.2, and Figure 5.3 were conducted by Nazanin Mohammadzadeh. Dr. Stewart and Sanche provided clinical samples used in Figure 5.4 and Table 5.2.

5.1 Abstract

APOBEC3 enzymes are a family host restriction factors that can restrict lentiviral replication. This family of cytidine deaminases can act as a barrier to viral infection by inducing mutagenesis of lentiviral genomes, such as HIV-1, through the deamination of cytosine forming promutagenic uracil in nascent single-stranded (-) DNA. HIV-1 suppresses APOBEC3 activity through the HIV-1 protein Vif that induces APOBEC3 degradation. Notably, for some APOBEC3s only some polymorphisms are active. Here we examined common polymorphisms of APOBEC3F that result in a protein with either 108A/231V or 108S/231I. We find that although both polymorphisms have HIV-1 restriction activity, APOBEC3F 108A/231V can restrict both Vif+ and Vif- HIV-1 more than APOBEC3F 108S/231I. The enhanced restriction resulted from enhanced encapsidation due to higher levels of steady state expression and incomplete Vif-mediated degradation. The APOBEC3F 108A/231V was not resistant to Vif-mediated degradation *per se*, but appeared to be maintained in cells at a level that exceeded the Vif degradation capacity. We show that individuals are commonly heterozygous for the APOBEC3F polymorphisms and in cell-based studies, these polymorphisms formed, independent of RNA, hetero-oligomers between each other and with APOBEC3G. Hetero-oligomerization with APOBEC3F 108A/231V resulted in partial stabilization of APOBEC3F 108S/231I and APOBEC3G in the presence of Vif. These data demonstrate how individual genotypes can affect HIV-1 restriction and are a functional example of how variations in host genotypes can protect a population or individual against virus infections.

5.2 Importance

APOBEC3 enzymes are a barrier to HIV-1 infection and suppress HIV-1 replication by inducing mutagenesis of HIV-1 proviral DNA. To counteract this host restriction system HIV-1 uses Vif to induce degradation of relevant APOBEC3 enzymes. Although there is a constant battle between host and virus for suppression of replication and a counteraction, the HIV-1 mutation rate far exceeds that of humans and facilitates a faster rate of adaptation. As a result, protective measures such as having multiple host genotypes can guard populations by slowing the rate of viral adaptation. We examined if two common genotypes of APOBEC3F were an example of multiple genotypes that can help protect the host. We found that the two common APOBEC3F genotypes most often occur as heterozygous alleles and that the one allele that was more protective could

stabilize APOBEC3F and APOBEC3G in cells to increase virus encapsidation and partially protect against Vif mediated degradation.

5.3 Project Summary

Here we assessed the A3F 108S/231I and A3F 108A/231V variants for their ability to restrict HIV-1 replication in the presence and absence of Vif. Based on past evidence that A3F can interact with itself and A3G, in an RNA-independent manner, we assessed how each A3F variant restricted HIV-1 replication when expressed together and in the presence of A3G (*33*). Since the A3F A108S/V231I are in strong linkage disequilibrium and have been identified to commonly occur together in multiple genetic studies, we refer to these variants as 231V or A3F 231I, since this variation was correlated epidemiologically in affecting HIV-1 progression (*89, 97, 98*). We find that the A3F 231V is a more efficient restriction factor for HIV-1 than A3F 231I. The increased restriction is due to an increase in A3F steady state levels in cells. This increased steady state level of A3F 231V resulted in a partial resistance to Vif. Genotyping of a donor population showed that the A3F 231V and 231I alleles most often occur together, rather than the homozygous alleles. This appeared to have a functional benefit since the interaction of A3F 231V could also increase the Vif resistance of A3G and A3F 231I, providing a reasoning for how high numbers of A3 induced mutations may be acquired in HIV-1+ individuals.

5.4 Introduction

5.4.1 APOBEC3 functional overview

APOBEC3 (A3) enzymes are a family of deoxycytidine deaminases that can act as host restriction factors for HIV-1 (*1, 131*). The A3 enzymes deaminate cytosine in single-stranded (ss) DNA to form promutagenic uracil (*3, 132-134*). In the right context, this uracil is used as an immune defense (*46, 91, 135-138*). Namely, of the seven human A3 enzymes, there are five that can restrict the replication of HIV-1 through the formation of uracils, A3D, A3F, A3G, A3H (haplotypes II, V, and VII), and A3C (S188I) (*45, 52, 139*). In an HIV-1 infected CD4+ T cell, these enzymes can become encapsidated into assembling virions through an interaction with cellular or HIV-1 genomic RNA that is also bound to HIV-1 Gag (*57, 64*). After virion maturation and infection of a target cell, the encapsidated A3 enzymes can access single-stranded HIV-1 (-) DNA after it is synthesized and exposed through reverse transcription and RNase H activity (*140*). The cytosines

that become deaminated to uracil are copied by the reverse transcriptase during (+) DNA synthesis and result in guanine to adenine (G→A) mutations on the coding strand (*3, 132-134, 140*). These mutations can induce DNA repair enzyme mediated degradation of proviral DNA or result in functional inactivation of integrated proviral DNA (*16*). HIV-1 Vif suppresses the encapsidation and activity of A3 enzymes by mediating their polyubiquitination and proteasomal degradation (*74, 137, 141-146*). HIV-1 Vif hijacks a Cullin 5 E3 ubiquitin ligase complex and acts as the substrate receptor by replacing the host protein SOCS2 (*74*). This enables HIV-1 Vif to directly interact with components of the ubiquitin ligase complex, Elongin C and Cullin 5 (*147*). HIV-1 Vif can also directly interact with A3D, A3F, A3G, A3H, and A3C, which induces their ubiquitination and proteasomal degradation. For stability, Vif also interacts with the transcription cofactor CBF-β (*75, 148*).

5.4.2 A3 and HIV-1 Vif Interactions

Despite the presence of Vif, A3 enzymes are still able to be encapsidated into HIV-1 virions and induce mutagenesis (*87, 149-159*). Although translation of Vif mRNA results in high levels of Vif in cells before virus particle assembly, which would deplete A3s from infected CD4+ T cells, the induced degradation of A3s is not complete (*160*). Evidence of this comes from high levels of A3 induced stop codons and missense mutations in HIV-1 proviral DNA in acute and chronic proviral DNA isolated from HIV-1+ individuals (*87*). Sequencing of integrated proviral genomes from multiple studies have revealed that the mutations are present on the (+) DNA in both a 5′GG→5′AG context and a 5′GA→5′AA context (*82, 87, 149-159*). The 5′GG context originates primarily from A3G that deaminates preferentially in 5′C<u>C</u>A motifs on the (-) DNA (underlined C deaminated) (*140*). This sequence context most often results in mutations that introduce stop codons and mutations at glycines (*4*). The 5′GA context can originate from A3D, A3F, A3H, and A3C (S188I) due to a preference to deaminate in 5′T<u>C</u> motifs on the (-) DNA (*46, 82, 136, 161*). Mutations at the 5′T<u>C</u> motif can also introduce stop codons and missense mutations at a larger variety of amino acids (*4*). Although A3G is the most active deaminase in a ΔVif HIV-1 infection, it is also very sensitive to Vif (*50, 162*). Thus, despite there being lower activity from A3D and A3F in ΔVif HIV-1 infections, there is some evidence that A3F is less sensitive to Vif, and as a result, may be a larger contributor to mutagenesis during HIV-1 infections (*45, 50, 97, 163, 164*). There is also evidence that in chronically infected individuals that HIV-1 Vif can acquire mutations that result in partial activity, thereby enabling A3 encapsidation and mutagenesis (*88*). However,

it is unlikely that these mutations in Vif are acquired early in infection and A3-mediated mutagenesis is found in acute infections suggesting that other mechanisms can enable A3 encapsidation in the presence of Vif (*87*).

5.4.3 Vif resistance observed in common A3 SNPs

Resistance to Vif mediated degradation has been characterized for some A3 enzymes. A3H exists as seven major haplotypes in humans with differential protein stability, HIV-1 restriction ability, and sensitivity to Vif (*90, 91, 165*). These haplotypes result from a combination of five different single-nucleotide polymorphisms (SNPs) (*90, 91, 165*). The diversity of A3H in the human population necessitates constant adaptation of HIV-1 Vif to A3H, especially since stable and restrictive A3H haplotypes (II, V, and VII) show population stratifications (*82, 90, 91*). This dynamic results in a situation where A3H can act as a transmission barrier for some HIV-1 strains or at least increase the time until CD4+ T cell counts decrease or until the development of AIDS (*81, 82*). The number of SNPs in other A3s is significantly lower and most commonly they have been found to result in inactivation of the A3 or no effect (*89, 96, 166*). Exceptions are A3D and A3C in which SNPs have resulted in increased activity (*52, 89, 139*). Interestingly, despite cell-based assays not uncovering differences in a common variant of A3F A108S/V231I, a population-based analysis of multiple pre-treatment cohorts of HIV-1+ individuals showed that the A3F 231V variant was associated with lower viral load and slower rate of progression to AIDS (*89, 97, 98*). The authors suggested that the A3F 231 position may increase resistance to Vif (*97*). Despite this amino acid not being within the canonical Vif binding interface on A3F, this observation was also previously reported for A3F 108A/231V (*136, 167-170*).

5.5 Materials and Methods

5.5.1 Donor PBMC and DNA isolation

PBMCs were isolated from whole blood samples using Ficoll-Paque (GE-Healthcare) density gradient centrifugation in Sepmate50 tubes (STEMCELL Technologies). DNA was isolated from donor PBMCs using a PAXgene Blood DNA Kit (Qiagen).

5.5.3 Sanger Sequencing genotyping

The A3G and A3F polymorphic sites were amplified using PCR. The A3G H186R polymorphism was detected using the following primers: (forward) 5′ AATTTCAGCACTGTTGGAGC and (reverse) 5′ GAGAATCTCCCCCAGCAT. The A3F V231I polymorphism was detected using the following primers: (forward) 5′ AGCCTATGGTCGGAACGAAA and (reverse) 5′ CTGGTTTCGGAAGACGCC. The PCR amplification used Q5 High Fidelity polymerase (New England Biolabs) and cycling conditions were 95° C for 5 min followed by 45 cycles of 95° C for 10 sec, 52° C for 20 sec, 72 ° C for 20 sec. Amplicons were cloned into pJET1.2 using the CloneJET PCR cloning kit (ThermoScientific). Amplicons were Sanger sequenced at the National Research Council of Canada (Saskatoon, Saskatchewan). The resulting A3 sequences were translated using the Expasy Translate tool to identify polymorphisms (*125*), clones from each donor were analyzed for each A3 condition and their respective polymorphisms to determine homozygosity or heterozygosity of each allele in combination with allelic frequency.

5.5.4 Plasmids and Transfection Conditions

To express two A3 enzymes on a single-cell level, the pVIVO2 plasmid (Invivogen) that contains two transcriptional units was used. For comparison to double expression experiments, pVIVO2 was also used to express single A3 enzymes. The cloning of A3G-HA, A3F 108S/231I-V5, and A3G-HA/A3F 108S/231I-V5 into pVIVO2 was previously described (*162*). The pcDNA A3F 108A/231V was kindly provided by Reuben Harris (University of Minnesota) and subcloned into pVIVO2 using the same cloning strategy as previously reported to create A3F 108A/231V-V5, A3F 108A/231V-HA, and A3G-HA/A3F 108A/231V-V5, and A3F 108A/231V-HA/A3F 108S/231I-V5 (*162*). Since the A3F polymorphisms at 108 and 231 are in strong linkage disequilibrium, for brevity we refer to the A3F forms as either 231I or 231V (*89, 97, 98*).

5.5.5 Single-cycle infectivity assay

VSV-G pseudotyped HIV-1 LAI ΔVif ΔEnv were produced by transfecting 1 x 10^5 293T cells per well of a 12-well plate. The 293T cells were maintained in DMEM with 10% FBS. GeneJuice (Novagen) transfection reagent was used as per manufacturer's protocol. Cells were transfected with 500 ng of pHIV-1 LAI ΔVif ΔEnv, 180 ng of pMDG, which expresses VSV-G, and 25, 50,

or 100 ng of A3 expression plasmid (124). The following pVIVO2 constructs were transfected for expression of tagged A3s: A3G-HA, A3F 231V-V5, A3F 231I-V5, A3G-HA/A3F 231I-V5, A3G-HA/A3F 231V-V5, and A3F 231V-HA/A3F 231I-V5. The following pcDNA3 constructs were transfected for expression of untagged A3s: A3G, A3F 231V, and A3F 231I. To equalize the amount of plasmid DNA transfected, empty pVIVO2 or pcDNA3 vectors were used. After 16 h the media was changed and virus containing supernatants were harvested 24 h after the media change. Supernatants were filtered through a 0.45-micron polyvinylidene difluoride (PVDF) syringe filter. For infection of a reporter cell line to determine infectivity 1 x 10^4 cells per well of a 96-well plate containing either HeLa CD4+ HIV-1 LTR-β-gal (MAGI) or TZM-Bl cells were infected with a serial dilution of virus in the presence of 8 μg/mL polybrene. The LTR-β-gal provides β-galactosidase under HIV-1 Tat promoter control. Forty-eight hours after infection the cells were washed with PBS and infectivity was measured through colorimetric detection using a β-galactosidase assay reagent (Pierce) and spectrophotometer. Infectivity of each virus was compared by using the infectivity of the No A3 condition as 100%.

5.5.6 Quantitative Immunoblotting

293T cells expressing No A3, A3G-HA, A3F 231V-V5, A3F 231I-V5, A3F-231V-HA, A3G-HA/A3F 231I-V5, A3G-HA/A3F 231V-V5, and A3F 231V-HA/A3F 231I-V5 from the single-cycle infectivity assays were detected using rabbit anti-HA (1:1000, Sigma) or rabbit anti-V5 (1:500, Sigma). 293T cells expressing No A3 or untagged A3G, A3F 231V, or A3F 231I were detected with ApoC17 rabbit antiserum (1:1000; Cat# 10082, NIH AIDS Reagent Program) for A3G or C-18 polyclonal rabbit antibody (1:500; GTX47211 (Gentex) for A3F (112). A3s were detected in cell lysates and virions. Cells were lysed using 2x Laemmli Buffer and 40 μg total protein was used. Virus was concentrated using Retro-X (Clontech) following the manufacturer's protocol and 20 μL of concentrated virus was used. Loading controls for cell lysates (α-tubulin, Sigma) and virus (p24, Cat #3537, NIH AIDS Reagent Program) were detected with mouse monoclonal antibodies. Secondary detection was performed using Licor IRDye antibodies produced in goat (IRDye 680-labeled anti-rabbit and IRDye 800-labeled anti mouse).

To detect Vif-mediated degradation of A3s, 1 x 10^5 293T cells per well of a 24-well plate were transfected with 50 ng of A3F 231I-V5, 10 ng A3F 231V-V5, or empty pcDNA3 (No A3) and a titration of Vif_{LAI} expression plasmid (0,10, 25, 50, 100, and 200 ng). A 5-fold greater amount of

A3F 231I was required to equalize the expression of both A3Fs. GeneJuice (Novagen) transfection reagent was used as described by the manufacturer. The media (DMEM and 10% FBS) was changed 16 h after the transfection. The cells were harvested 24 h after the media change with 2x Laemmli buffer and 40 μg total protein was used. The V5-tagged A3F and α-tubulin was detected as described above.

Cell lysates were quantitatively analyzed where indicated in figures. Quantitation of band intensities was performed using Odyssey Software with normalization to either α-tubulin or p24, which were detected in parallel on the same blot with experimental samples. Relative expression levels were then determined by comparison to a control sample set at 1.

5.5.7 Co-immunoprecipitation

The co-immunoprecipitation protocol was from Baig *et al.* 2014 (*162*)**.** In brief, 2.5 x 10^6 293T cells per 75 cm^2 flask were transfected with 1 μg total DNA using GeneJuice (Novagen) as per manufacturer instructions. The plasmids transfected expressed A3G-HA, A3F 231I-V5, A3G-HA/A3F 231I-V5, A3G-HA/A3F 231V-V5 or A3F 231V-HA/A3F-231I-V5. Sixteen hours post transfection media was changed with fresh DMEM+10% FBS. Forty-hours post transfection, cells were washed with PBS and lysed with buffer (50 mM Tris pH 7.4, 1% nonident-p40, 0.1% sodium deoxycholate, 150 mM NaCl, protease inhibitor (Roche)). Clarified supernatants were split into two fractions that were either used in a co-immunoprecipitation with nonspecific mouse IgG (mock) or the mouse anti-V5 (experiment). The co-immunoprecipitations were conducted in the presence of RNase A (*33*). A mouse anti-V5 (1:1000, Sigma) was used for immunoprecipitations and proteins were detected on the membrane using the same V5 antibody, mouse anti-HA (1:1000, Sigma), and polyclonal rabbit anti-α-tubulin (1:1000; Sigma). Secondary antibodies used were Licor IRDye antibodies IRDye 800-labled goat anti-mouse and 680-labled goat anti-rabbit secondary antibody.

5.5.8 Proviral Sequencing

For proviral sequencing, 1 x 10^5 293T cells were infected with 40 μL of supernatant containing virus in the presence of 8 μg/mL polybrene. The plates were spinoculated at 800 x *g* for 1 h. Cells were harvested after 48 h by removing the media, washing with PBS, and lysing the cells and extracting DNA with DNAzol (Invitrogen). The PCR amplification of the protease region of HIV-

1 (351 bp) was conducted as previously described Ara *et al.* 2014 (*50*). Sequences were analyzed with Clustal Omega and Hypermut software programs (*126*) (*127*).

Ethics Statement.

This study has been reviewed and approved by the Research and Ethics Board at the University of Saskatchewan (Bio# 14-62 and 13-293). Peripheral Blood Mononuclear Cells (PBMCs) were isolated from de-identified individuals Donors were either HIV+ and recruited through the Positive Living Program (PLP) at the Royal University Hospital or West Side Clinic in Saskatoon or HIV- Donors (HD).

5.6 Results

5.6.1 A3F 231V more efficiently restricts HIV-1 infection than A3F 231I

The A3F 231V allele has been shown to be protective in HIV-1+ individuals, but the reason for this was not fully explored (*97*). Here we conducted an assessment of the HIV-1 restriction ability of A3F 231V and A3F 231I. The effect of untagged A3F 231V and A3F 231I on VSV-G pseudotyped ΔVif HIV-1 infectivity was assessed using transient transfection of different amounts of A3F vector in single-cycle infectivity assays. Over a range of A3F plasmid transfection amounts, the A3F 231V restricted ΔVif HIV-1 1.4-fold (25 ng, $p \leq 0.05$), 1.6-fold (50 ng, $p \leq 0.01$), and 4.0-fold (100 ng, $p \leq 0.001$) more than A3F 231I (Figure 5.1A). Overall the ability for A3F to restrict ΔVif HIV-1 was still 1.5-fold (25 ng, $p \leq 0.01$), 1.8-fold (50 ng, $p \leq 0.01$), and 13.6-fold (100 ng, $p \leq 0.001$) less than A3G, for the more active A3F 231V variant, consistent with previous reports (Figure 5.1A) (*45*). Also consistent with A3F 231V being protective in HIV-1+ individuals, the A3F 231V enabled some restriction of virus replication in the presence of Vif, which was not observed for A3F 231I or A3G (Figure 5.1B, $p \leq 0.01$ for 50 ng and 100 ng).

To confirm that decreases in infectivity in the presence and absence of Vif correlated with virion encapsidation of the A3 enzymes, the cell lysates and virions were analyzed after proteins were resolved by SDS-PAGE and transferred to membranes for immunoblotting. The native antibody

used to detect A3F was only able to detect cellular expression and not virion encapsidation, except at the 100 ng transfection amount. These blots showed that the A3F 231V had a higher steady state expression level than A3F 231I in the absence of Vif (5- to 10- fold) and presence of Vif (12.5-fold, Figure 1C-E), consistent with higher decreases in HIV-1 infectivity in both of these conditions (Figure 5.1A-B). At the highest plasmid transfection amount (100 ng), the A3F variants could be detected in the virus particles. The A3F 231V was encapsidated 5-fold more than A3F 231I in the absence of Vif and in the presence of Vif encapsidation of A3F231I was not detectable (Figure 5.1E). These data suggest that the A3F 231V is partially resistant to Vif mediated degradation. Similar limitations with native antibody were identified with A3G in that the cellular expression levels were easily detected with the native antibody, but the encapsidation was not detected until the plasmid transfection amount reached 100 ng, despite evidence of HIV-1 restriction in the single cycle infectivity assays (Figure 5.1A and F-H). Although significant amounts of A3G were detected in the virus in the presence of Vif at the higher transfection amount (Figure 5.1H), there was no observable decrease in infectivity (Figure 5.1B), presumably due to Vif-mediated inhibition of A3G catalytic activity and processivity (*171, 172*).

To confirm the partial resistance to Vif-mediated degradation by A3F 231V we used V5-tagged A3F 231V and A3F 231I constructs, which would facilitate quantification of immunoblots at lower plasmid transfection amounts. We also tested if the V5 tag would effect differences in infectivity using a single cycle infectivity assay. In the absence of Vif, the A3F 231V still restricted HIV-1 more than the A3F 231I by 1.4-fold (50 ng) or 2-fold (100 ng), consistent with the untagged enzymes (Figure 5.1A and Figure 5.2 A). The V5 tag had no significant effect on the A3F231V restriction ability, although there appeared to be some enhancement the A3F 231I restriction ability in the absence of Vif (Figure 5.1A and Figure 5.2A). Nonetheless, comparable differences in restriction were found with and without the tag at the 50 ng plasmid transfection amount and was still observable at the 100 ng transfection amount (Figure 5.1A and Figure 5.2A). Further, in the presence of Vif, the A3F 231V was still able to restrict the HIV-1 infectivity more than A3F 231I (Figure 5.1, 100 ng). Altogether, these results indicated that a tag could be used to facilitate immunoblotting.

The immunoblots of the V5-tagged A3F showed that at all transfection amounts the A3F 231V had higher steady state expression levels in cells, suggesting that the variant was more stable. The

A3F 231V had steady state levels that were 5-fold more than A3F 231I (Figure 5.2C-E), consistent with the untagged enzymes (Figure 5.1C-E). The A3F 231V was also present in cell lysates in the presence of Vif at all plasmid transfection amounts, in contrast to A3F 231I that could primarily be detected when 100 ng plasmid was transfected (Figure 5.2C-E). In the virus particles and the absence of Vif, the A3F 231V was encapsidated at 3- to 10- fold higher levels than A3F 231I, consistent with higher cellular expression levels (Figure 5.1C-E and Figure 5.2C-E). In the presence of Vif, the A3F 231I encapsidation was still not detectable at any transfection amount (Figure 5.2C-E). These data indicated that either the A3F 231V was more resistant to Vif or the steady-state expression levels in cells exceeded the ability of Vif to induce its degradation.

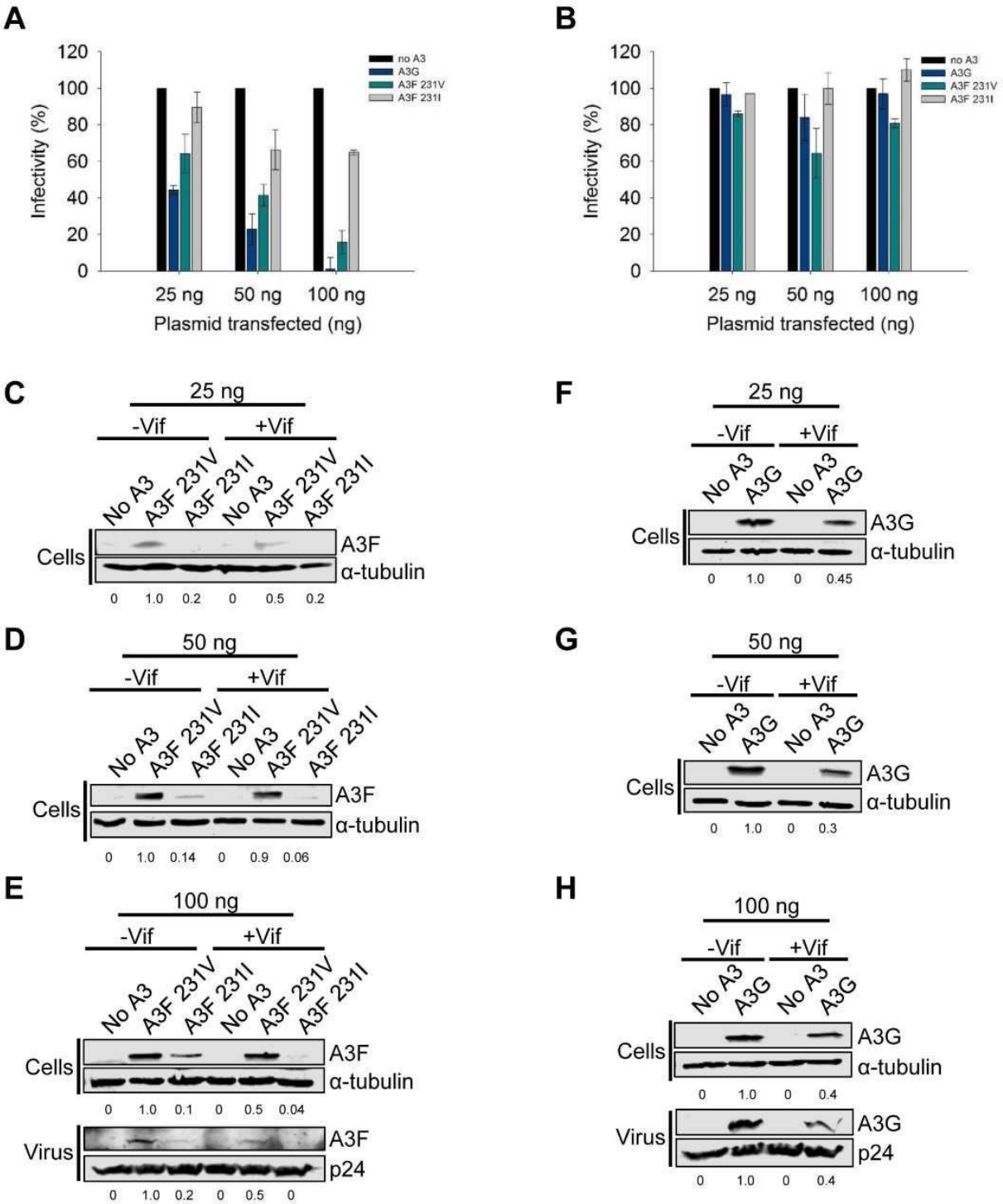

A
Infectivity (%)
120
100
80
60
40
20
0
no A3
A3G
A3F 231V
A3F 231I
25 ng
50 ng
100 ng
Plasmid transfected (ng)
B
Infectivity (%)
120
100
80
60
40
20
0
no A3
A3G
A3F 231V
A3F 231I
25 ng
50 ng
100 ng
Plasmid transfected (ng)
C
25 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
A3F
α-tubulin
0 1.0 0.2 0 0.5 0.2
D
50 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
A3F
α-tubulin
0 1.0 0.14 0 0.9 0.06
E
100 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
A3F
α-tubulin
0 1.0 0.1 0 0.5 0.04
Virus
A3F
p24
0 1.0 0.2 0 0.5 0
F
25 ng
-Vif
+Vif
No A3
A3G
No A3
A3G
Cells
A3G
α-tubulin
0 1.0 0 0.45
G
50 ng
-Vif
+Vif
No A3
A3G
No A3
A3G
Cells
A3G
α-tubulin
0 1.0 0 0.3
H
100 ng
-Vif
+Vif
No A3
A3G
No A3
A3G
Cells
A3G
α-tubulin
0 1.0 0 0.4
Virus
A3G
p24
0 1.0 0 0.4

Figure 5. 1 A3F 231V and A3F 231I have different HIV-1 restriction abilities

(A) HIV-1 ΔVif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 ΔVif that was produced in the absence or presence of untagged A3G, A3F 231V or A3F 231I. Results were normalized to the no A3 condition. (B) HIV-1 +Vif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 +Vif that was produced in the absence or presence of untagged A3G, A3F231V or A3F231I. Results were normalized to the no A3 condition. (A-B) Error bars represent the standard deviations of the mean calculated from three independent experiments. Immunoblotting with A3F native antibody was used to detect (C) 25 ng (D) 50 ng and (E) 100 ng transfected A3F expressed in cells and encapsidated into HIV-1 virions in -Vif and +Vif conditions. Immunoblotting with A3G native antibody was used to detect (F) 25 ng (G) 50 ng and (H) 100 ng transfected A3G expressed in cells and encapsidated into HIV-1 virions in -Vif and +Vif conditions. The cell lysate and virion loading controls were α-tubulin and p24, respectively. (C-H) Expression levels shown below blots were calculated by setting a –Vif condition to 1.0 and determining the relative values of other lanes.

5.6.2 A3F 231V is more protected from Vif mediated degradation than A3F 231I due to higher steady state protein levels

To differentiate between these possibilities regarding A3F 231V partial resistance to Vif mediated degradation, we expressed increasing amounts of Vif from an expression vector in the presence of 10 ng of transfected A3F 231V-V5 and 50 ng A3F 231I-V5. Quantification of the immunoblot demonstrated that at these transfection amounts, the steady state levels of both A3F 231V and A3F 231I were similar (Figure 5.3A). With increasing amounts of transfected Vif, the induced degradation of both A3F variants were similar. Thus, the resistance to Vif was not an inherent biochemical property of the A3F 231V (Figure 5.3A-B). Rather the data suggest that the partial resistance to Vif mediated degradation observed in single-cycle infectivity assay blots was due to the higher steady state levels of A3F 231V that exceeded Vif degradation capacity (Figure 5.1C-E and Figure 5.2C-E).

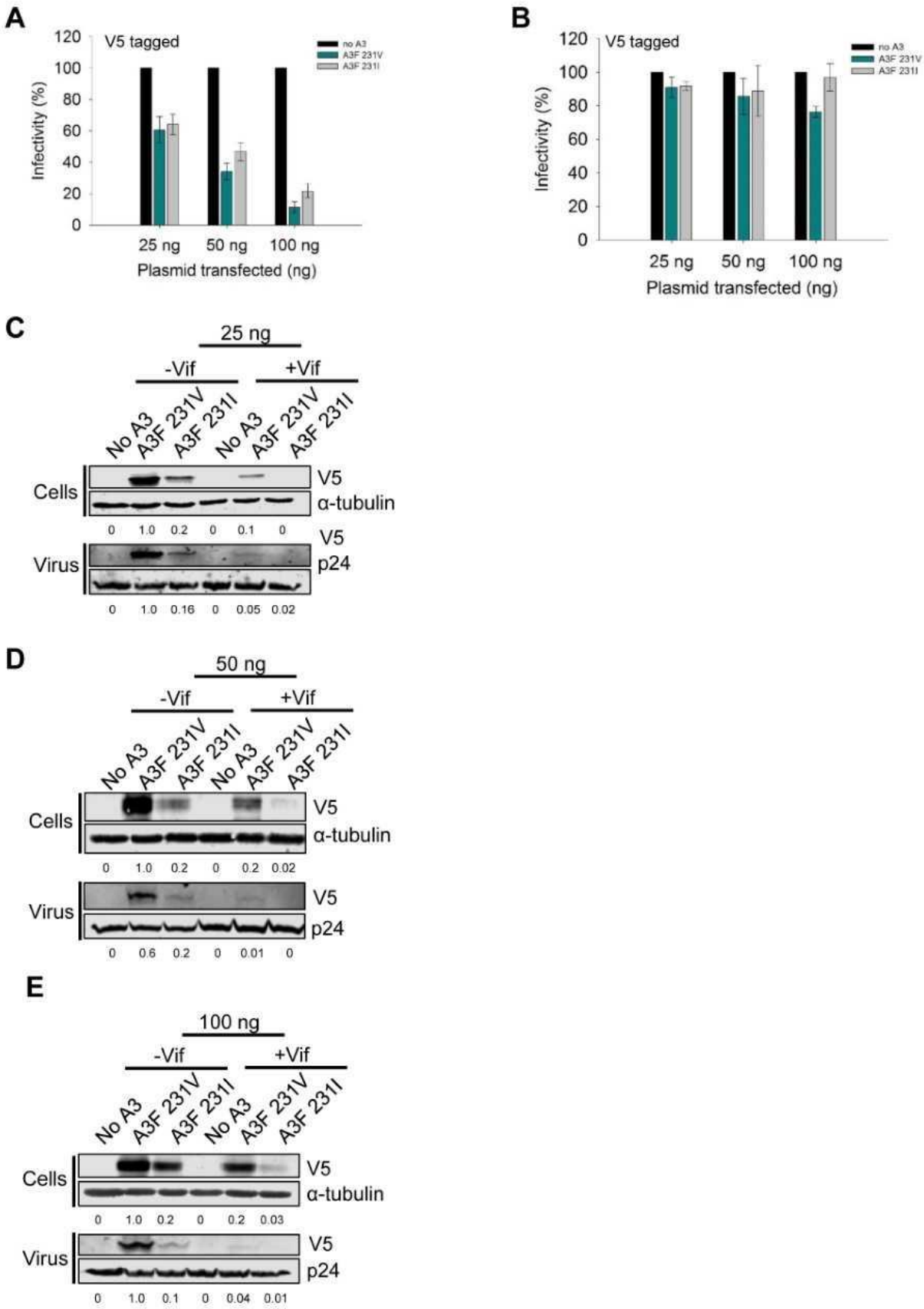

A
V5 tagged
no A3
A3F 231V
A3F 231I
Infectivity (%)
120
100
80
60
40
20
0
25 ng
50 ng
100 ng
Plasmid transfected (ng)
B
V5 tagged
no A3
A3F 231V
A3F 231I
Infectivity (%)
120
100
80
60
40
20
0
25 ng
50 ng
100 ng
Plasmid transfected (ng)
C
25 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
V5
α-tubulin
0 1.0 0.2 0 0.1 0
Virus
V5
p24
0 1.0 0.16 0 0.05 0.02
D
50 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
V5
α-tubulin
0 1.0 0.2 0 0.2 0.02
Virus
V5
p24
0 0.6 0.2 0 0.01 0
E
100 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
V5
α-tubulin
0 1.0 0.2 0 0.2 0.03
Virus
V5
p24
0 1.0 0.1 0 0.04 0.01

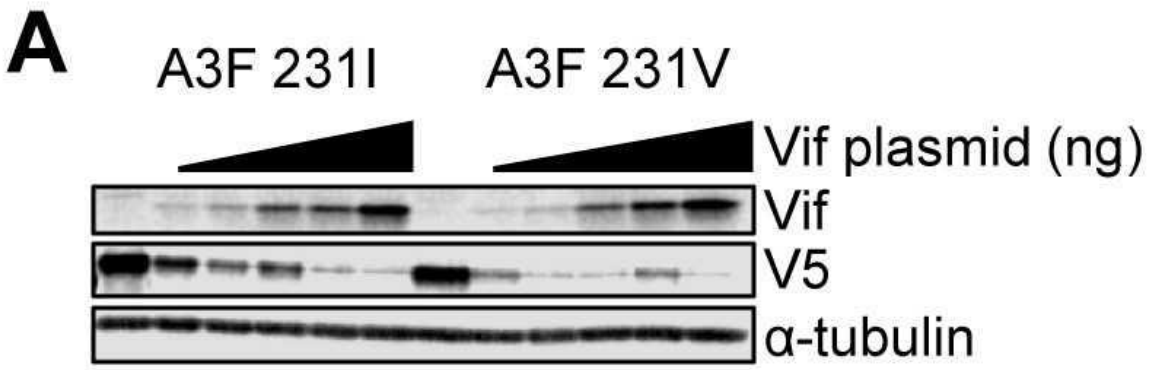
Figure 5. 2 A3F 231V is expressed at higher steady state levels in cells

(A) HIV-1 ΔVif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 ΔVif that was produced in the absence or presence of A3F 231V-V5 or A3F 231I-V5. Results were normalized to the no A3 condition. (B) HIV-1 +Vif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 +Vif that was produced in the absence or presence of A3F 231I-V5 or A3F 231I-V5. Results were normalized to the no A3 condition. (A-B) Error bars represent the standard deviations of the mean calculated from three independent experiments. Immunoblotting to detect the V5 tag was was carried out for (C) 25 ng (D) 50 ng and (E) 100 ng transfected A3F to detect the amount of A3F 231V and A3F 231I expressed in cells and encapsidated into HIV-1 virions in –Vif and +Vif conditions. The cell lysate and virion loading controls were α-tubulin and p24, respectively. (C-E) Expression levels shown below blots were calculated by setting a –Vif condition to 1.0 and determining the relative values of other lanes.

5.6.3 A3F 231V more efficiently restricts HIV-1 infection through a deamination dependent mechanism

A3F has been shown to inhibit HIV-1 through deamination dependent and independent mechanisms (*32, 33, 136-138, 161, 173-175*). To determine if the enhanced restriction of A3F 231V was due to enhanced deamination, we PCR amplified a 351 bp region of HIV-1 *protease* for sequencing under the conditions shown in Figure 1A (untagged A3s, 50 ng transfection amount). We observed in the absence of Vif that G→A mutations per kb were 2.46 for A3F 231V and 1.42 for A3F 231I (Table 5.1). This ~2-fold difference in mutations is consistent with the decrease in infectivity (Figure 1A, 50 ng). However, both A3F variants were still less able to induce mutations than A3G (Table 5.1, 13.4 G→A mutations/kb). This at least 5-fold difference in A3F and A3G induced mutations is consistent with previous data and has been found to be due to differences in enzyme processivity (*50*). As a result, these data emphasize the importance of the inherent biochemical characteristics in restriction capacity between A3s. However, when comparing A3F variants, the level of encapsidation appears to be a determining factor that can increase A3F restriction efficiency for HIV-1 (Table 5.1, Figure 5.1E, and Figure 5.2C-E).

Importantly in the presence of Vif, the number of A3-induced mutations by A3F 231V and A3G were less disparate. A3G did not show a decrease in infectivity in the presence of Vif and accordingly induced 10-fold less mutations than in the absence of Vif (Table 5.1, 1.3 G→A mutations/kb). Consistent with the ability to partially avoid Vif-mediated degradation, the A3F 231V only had a 4.8-fold decrease in induced mutations (Table 5.1, 0.51 G→A mutations/kb). Thus, in the presence of Vif the difference between the A3F 231V and A3G was only 2.5-fold. In the presence of Vif, the A3F 231I induced mutations decreased 11-fold to 0.13 G→A mutations/kb, resulting in 10-fold less mutations than A3G and 4-fold less mutations than A3F 231V (Table 5.1).


A
A3F 231I
A3F 231V
Vif plasmid (ng)
Vif
V5
α-tubulin

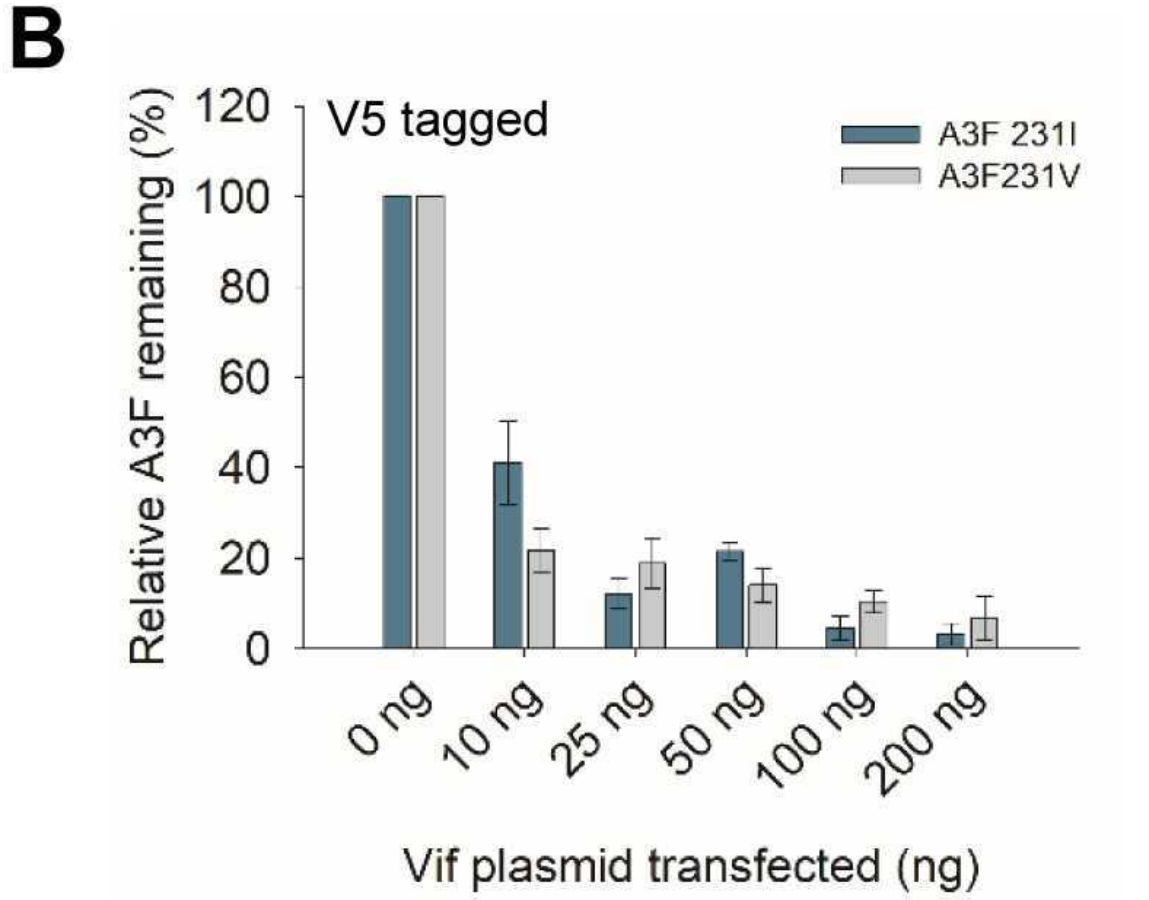

B
V5 tagged
A3F 231I
A3F231V
Relative A3F remaining (%)
120
100
80
60
40
20
0
0 ng
10 ng
25 ng
50 ng
100 ng
200 ng
Vif plasmid transfected (ng)

Figure 5. 3 A3F 231V and A3F 231I are equally sensitive to Vif-mediated degradation

(A) A3F 231V-V5 and A3F 231I-V5 plasmids were transfected into 293T cells with amounts that resulted in equal steady state expression levels. Increasing amounts of Vif were cotransfected. The amount of Vif-mediated degradation was detected by analysing the intensity of the bands detected with antibody to the V5 tag. (B) Quantified intensities of the bands from immunoblotting are shown. Error bars represent the standard deviations of the mean calculated from three independent experiments.

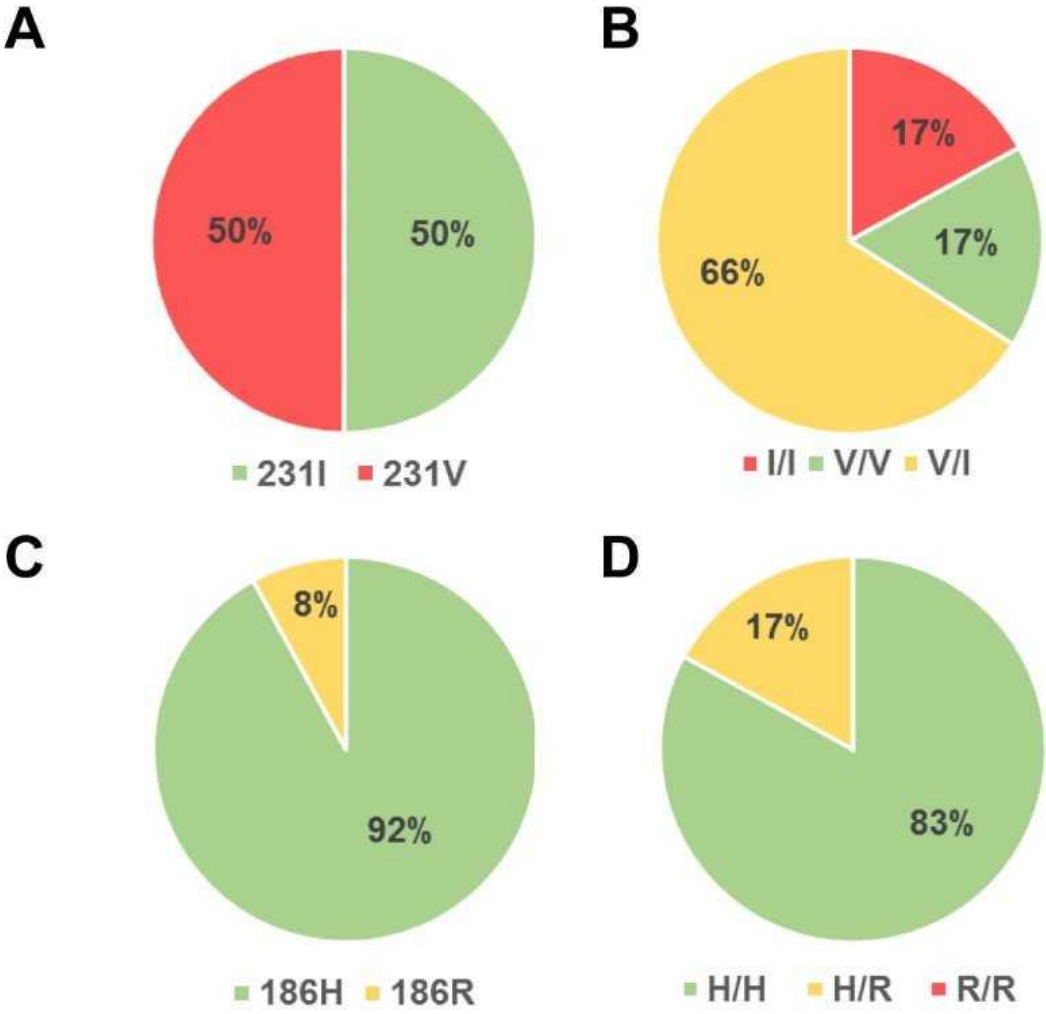
Table 5. 1 Analysis of A3-induced mutagenesis in HIV-1 proviral DNA

Virus Condition	A3 enzyme	Base pairs sequenced	G→A Mutations (per kb)
-Vif	No A3	7722	0.26
	A3F 231V	7722	2.46
	A3F 231I	7020	1.42
	A3G	6669	13.4
+Vif	No A3	7722	0.00
	A3F 231V	15795	0.51
	A3F 231I	15093	0.13
	A3G	20874	1.30

5.6.4 231V and A3F 231I hetero-oligomerize with each other and A3G

Based on these results showing that different A3F variants can be coexpressed within an individual, we determined whether the A3F variants could hetero-oligomerize with each other or A3G. A3F has been shown to form homo-oligomers and to hetero-oligomerize with A3G, in the absence of RNA (33). To determine oligomerization, we used a cotransfection scheme where one

A3 had an HA-tag and the other A3 had a V5-tag and determined if one A3 could co-immunoprecipitate the other A3. To ensure co-expression on a per cell basis we used the pVIVO2 plasmid that has two transcription units within a single vector. We used the V5-tagged A3 to immunoprecipitate the binding partner, in the presence of RNase A, and found through blotting for HA that A3G-HA can interact with A3F 231I-V5, as shown previously (Figure 5.4E) (33). We also determined that A3G-HA can interact with A3F 231V-V5 and that A3F 231V-HA can interact with A3F 231-V5 (Figure 5.4E).


A
50%
50%
231I
231V
B
17%
17%
66%
I/I
V/V
V/I
C
8%
92%
186H
186R
D
17%
83%
H/H
H/R
R/R

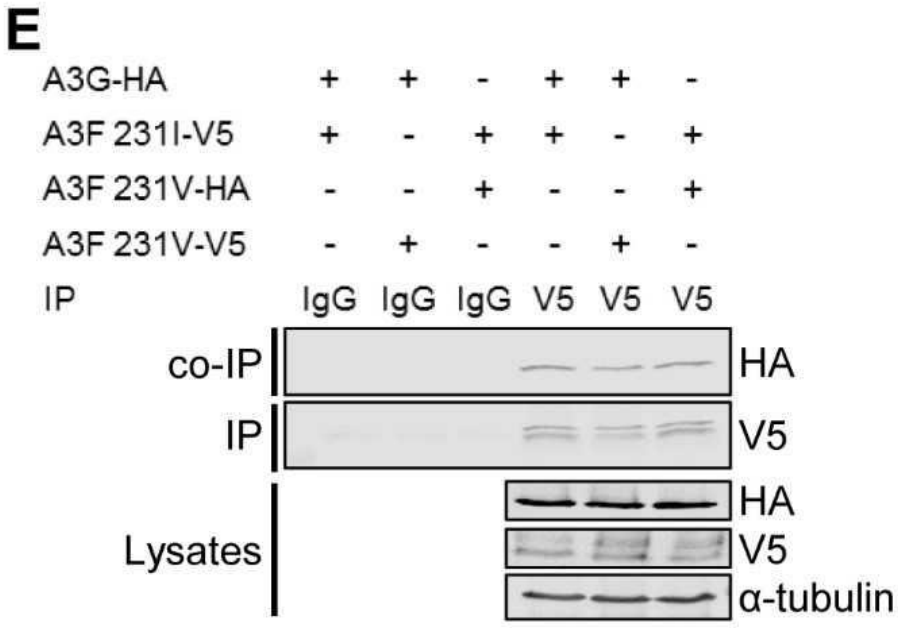

E
A3G-HA + + - + + -
A3F 231I-V5 + - + + - +
A3F 231V-HA - - + - - +
A3F 231V-V5 - + - - + -
IP IgG IgG IgG V5 V5 V5
co-IP
IP
Lysates
HA
V5
HA
V5
α-tubulin

Figure 5. 4 A3F 231V and A3F 231I commonly occur as a heterozygous genotype and can interact in cells

The (A) allelic frequency and (B) genotype of 6 donors was determined for A3F 231V and A3F 231I. The (C) allelic frequency and (D) genotype of 6 donors was determined for A3G 186H and A3G 186R. (E) The immunoprecipitation from cell lysates used either anti-V5 antibody or rabbit IgG (mock). Co-immunoprecipitation of A3F 231V-HA with A3F 231I-V5 or A3G-HA with A3F 231V-V5 and A3F 231I-V5 was detected through the HA tag. The presence of the V5 antibody in the co-immunoprecipitation samples is also shown. The experimental and mock samples were from the same lysate. The lysate blot demonstrates the cellular expression of α-tubulin, HA, and V5.

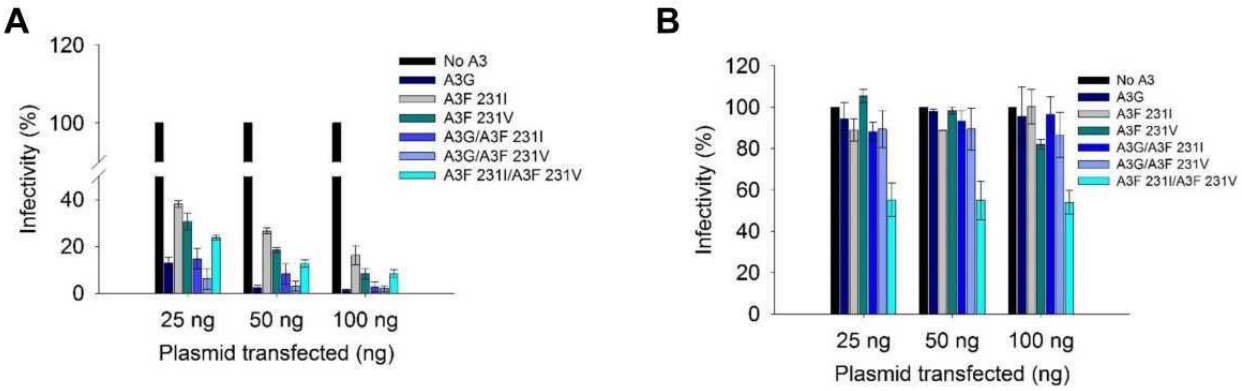
Table 5. 2 Genotype analysis of donors

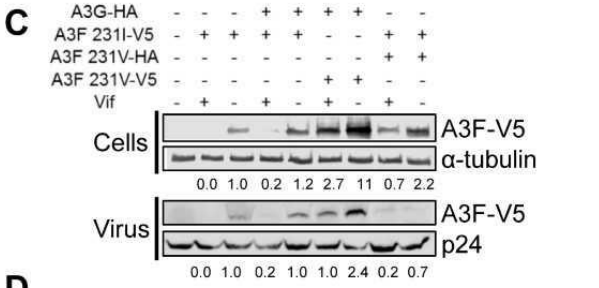
Donor	A3F 231VI	A3G H186R
PLP-002	V/V	H/H
PLP-004	I/V	H/H
PLP-007	I/I	H/R
HP02	I/V	H/H
HP03	I/V	H/H
HP06	I/V	H/H

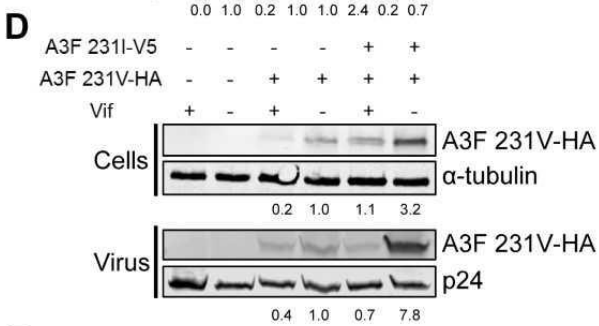
5.6.5 The presence of Vif A3F 231V promotes in higher steady state levels of A3F 231I and A3G

Using the same pVIVO2 expression vectors that were used in the co-immunoprecipitations we conducted single-cycle infectivity assays to assess if combined expression changed the HIV-1 restriction capability of the coexpressing A3s compared to the individual enzymes. In the absence of Vif, there was an increasing ability of the A3s to restrict HIV-1 with increasing amounts of A3 plasmid transfected (Figure 5.5A). With each incremental increase of transfected A3 plasmid (25ng, 50ng, 100ng) we observed a uniform decrease in percent infectivity, for simplicity of discussion we have chosen to 50ng condition to represent the experiment. Notably, the V5-tagged versions of A3F 231I and A3F 231V were able to decrease HIV-1 infectivity more than the untagged proteins (Figure 5.1B and 5.5A), but the difference in restriction between the two A3F variants was the same (Figure 5.1B, 50 ng, 1.6-fold, $p \leq 0.01$; Fig. 5.5A, 50 ng,1.5-fold, $p \leq 0.001$). Conversely, the coexpression of A3F 231V and A3F 231I (A3F 231V/A3F 231I) resulted in a significant 2-fold decrease in HIV-1 infectivity in comparison to A3F 231I and A3F 231V alone (Figure 5.5A, 50 ng, ($p \leq 0.001$) ($p \leq 0.05$). In the presence of Vif, the greater ability to restrict HIV-1 was most strikingly observed for coexpressed A3F 231V/A3F 231I, whereby a 1.7-fold decline in infectivity was observed in comparison to A3F 231I (Figure 5.5B, 50ng ($p \leq 0.01$). Immunoblotting demonstrated that the A3F 231V was not only more stable in cells, but that it resulted in increased steady state levels of coexpressed proteins, A3G or A3F 231I. In both the presence and absence of Vif, increased HIV-1 restriction correlated with increased cell expression of A3F 231I when coexpressed with A3F 231V (Figure 5.5C). In the absence of Vif and the presence of A3F 231V, the A3F 231I expression increased 2.2-fold (Figure 5C). In the presence of Vif and A3F 231V, the A3F 231I expression increased to a detectable level (Figure 5.5C). However, A3F 231V only increased A3F 231I encapsidation in the presence of Vif (Figure 5C). For A3F 231V, coexpression with A3F 231I increased A3F 231V cell expression 5-fold (+Vif) to 3-fold (-Vif) and resulted in 1.5-fold (+Vif) to 7.8-fold (-Vif) increases in encapsidation (Figure 5.5D). Altogether these data suggest that the enhanced restriction is primarily due to more A3F 231V being encapsidated.

Effects on encapsidation were also observed with A3G when coexpressed with A3F 231V (A3G/A3F231V). Consistent with previous data, co-expression of the A3G and A3F 231I resulted in 3- to 14-fold higher levels of ΔVif HIV-1 restriction than A3F 231 alone, depending on the transfection conditions. There was no enhancement of A3F 231I or A3G encapsidation levels in the absence of Vif, consistent with a previous report (*33*), but in the presence of Vif A3G was encapsidated 2-fold more when coexpressed with A3F 231I (Figure 5.5C and Figure 5.5E) (*33*). When comparing the infectivity variation between A3G and A3G/A3F 231V there was no significant difference observed (Figure 5.5, 50ng). However, when infectivity is compared between A3F 231V and A3G/A3F 231V a significant difference is observed, with the coexpressed A3G/A3F 231V exhibiting a 6-fold decline in infectivity in comparison to A3F 231V alone (Figure 5.5A, 50ng, ($p \leq 0.001$). Although not as prominent, a similar trend is also observed in the comparison of infectivity between A3G and A3G/A3F 231I, where again we see no significant difference in infectivity (Figure 5.5A, 50ng). But when A3F 231I is compared to A3G/A3F 231I we do see a significant difference, with A3G/A3F 231I exhibiting a 3-fold decline in percent infectivity in comparison to A3F I231 alone (Figure 5.5A, 50ng ($p \leq 0.001$). Consistent with these data, the virion encapsidation of A3G when coexpressed with A3F 231V was increased in the absence of Vif by 2-fold above the A3G/A3F 231I condition (Figure 5.5E). Altogether, the data indicate that when A3F 231V is expressed, it can increase steady state levels of A3F 231I and increase steady state levels and encapsidation of A3G (Figure 5.5C and Figure 5.5E). Consistent with the partial restriction ability of A3F 231V/A3F 231I in the presence of Vif (Figure 5.5B, 55% infectivity) there was encapsidation of A3F 231V and A3F 231I into virions even in the presence of Vif (Figure 5.5C and Figure 5.5D). Although A3G encapsidation in the presence and absence of Vif was increased when coexpressed with A3F 231V, the infectivity the presence of Vif was not decreased more than A3G alone, suggesting that Vif inhibited encapsidated A3G activity (*171*, *172*).


A
Infectivity (%)
120
100
40
20
0
No A3
A3G
A3F 231I
A3F 231V
A3G/A3F 231I
A3G/A3F 231V
A3F 231I/A3F 231V
25 ng
50 ng
100 ng
Plasmid transfected (ng)
B
Infectivity (%)
120
100
80
60
40
20
0
No A3
A3G
A3F 231I
A3F 231V
A3G/A3F 231I
A3G/A3F 231V
A3F 231I/A3F 231V
25 ng
50 ng
100 ng
Plasmid transfected (ng)


C
A3G-HA
A3F 231I-V5
A3F 231V-HA
A3F 231V-V5
Vif
Cells
A3F-V5
α-tubulin
0.0 1.0 0.2 1.2 2.7 11 0.7 2.2
Virus
A3F-V5
p24
0.0 1.0 0.2 1.0 1.0 2.4 0.2 0.7


D
A3F 231I-V5
A3F 231V-HA
Vif
Cells
A3F 231V-HA
α-tubulin
0.2 1.0 1.1 3.2
Virus
A3F 231V-HA
p24
0.4 1.0 0.7 7.8

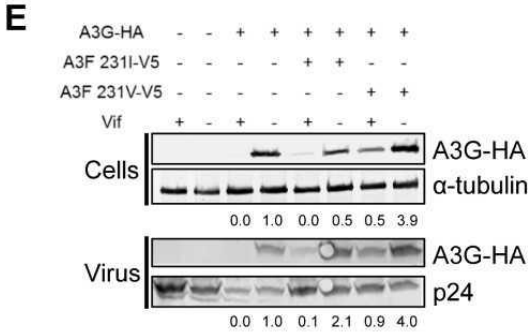

E
A3G-HA
A3F 231I-V5
A3F 231V-V5
Vif
Cells
A3G-HA
α-tubulin
0.0 1.0 0.0 0.5 0.5 3.9
Virus
A3G-HA
p24
0.0 1.0 0.1 2.1 0.9 4.0

Figure 5. 5 Co-Expression of A3F 231V and A3F 231I results in enhanced HIV-1 restriction ability

(A) HIV-1 ΔVif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 ΔVif that was produced in the absence or presence of A3G-HA, A3F 231V-V5, A3F 231V-HA, or A3F 231I-V5 in the combinations shown in the figure legend. (B) HIV-1 +Vif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 +Vif that was produced in the absence or presence of A3G-HA, A3F 231V-V5, A3F 231V-HA, or A3F 231I-V5 in the combinations shown in the figure legend. Results were normalized to the no A3 condition. (A-B) Error bars represent the standard deviations of the mean calculated from three independent experiments. (C) Immunoblotting with V5 antibody was used to detect cellular expression and virus encapsidation of A3F 231I-V5 and A3F 231V-V5 in the presence of A3G-HA. (D) Immunoblotting with HA antibody was used to detect cellular expression and virus encapsidation of A3F 231V-HA in the presence of A3F 231I-V5. (E) Immunoblotting with HA antibody was used to detect cellular expression and virus encapsidation of A3G-HA in the presence of A3F 231I-V5 and A3F 231V-V5. (C-E) Expression levels shown below blots were calculated by setting a – Vif condition to 1.0 and determining the relative values of other lanes.

5.6.6 In the presence and absence of Vif, coexpressed A3F 231V and A3F 231I induce more mutations in HIV-1 proviral DNA

To determine if the increased encapsidation of A3 enzymes resulted in increases in A3-induced mutations, we PCR amplified and sequenced a 351 bp portion of HIV-1 *protease*. The data demonstrated that the tagged versions of A3F 231V, A3F 231I, and A3G induced similar numbers of mutations to their untagged versions (+Vif and –Vif, Table 5.1 and Table 5.3). When the tagged A3F 231V and A3F 231I are coexpressed and encapsidate into ΔVif HIV-1 virons, there is a 2-fold increase in mutations from A3F 231I alone and a 1.4-fold increase in mutations for A3F 231V alone (Table 5.3). Based on analysis of encapsidation, this is likely due to increased encapsidation of A3F 231V (Figure 5.5D, 7.8-fold more encapsidated when coexpressed). The conditions where A3G was coexpressed with an A3F, in HIV-1 ΔVif producer cells, did not have significant increases in the numbers of mutations above A3G alone, but when we examined the ratio of A3G induced mutations to A3F induced mutations based on the mutational footprint, we observed a 2.6-fold (A3F 231V) and 2-fold (A3F 231I) increase in A3F-induced mutations in the presence of A3G (Table 5.3). Of more significance is that in the presence of Vif, there is a 2-fold increase in mutations from the A3F 231V alone condition (Table 5.3). Since A3F 231V and A3F 231I have the same deamination motif, we could not discern which variant is more active. However, based on analysis of encapsidation, since both had higher levels in virions, the deaminations could be due to both A3F variants.

Table 5. 3 Analysis of combined A3 expression on the induced mutagenesis in HIV-1 proviral DNA

Virus Condition	A3 enzyme	Base pairs sequenced	G→A mutations (total)	GG→AG mutations (total)	GA→AA mutations (total)	G→A mutations (per kb)	GG→AG mutations (per kb)	GA→AA mutations (per kb)
-Vif	No A3	10881	0	0	0	0	0	0
	A3F 231V-V5	10530	27	1	26	2.56	0.09	2.47
	A3F 231I-V5	8073	15	1	14	1.86	0.12	1.73
	A3F 231V-HA/A3F231I-V5	12285	45	1	44	3.66	0.08	3.58
	A3G	6669	121	109	11	18.1	16.3	1.64
	A3F 231V-V5/A3G-HA	10530	172	102	70	16.3	9.69	6.64
	A3F 231I-V5/A3G-HA	12987	134	95	39	10.3	7.32	3.00
+Vif	No A3	11934	0	0	0	0	0	0
	A3F 231V-V5	10530	6	1	5	0.57	0.09	0.47
	A3F 231I-V5	10881	1	0	1	0.09	0	0.09
	A3F 231V-HA/A3F231I-V5	11232	11	0	11	0.98	0	0.98

5.7 Discussion

In HIV-1 genomes isolated from HIV-1+ individuals, there are mutational footprints in the (+) DNA of both 5′GG→5′AG and 5′AG→5′AA motifs (*82, 87, 149-159*). The 5′GG→5′AG are clearly attributable to A3G (*140*). The origins of the 5′AG→5′AA mutations has been attributed to stable A3H haplotype (II, V, VII) induced mutations, due to an inability of some HIV-1 Vif variants to efficiently induce degradation of A3H (*81, 82*). The role of A3D in inducing these mutations is known to be minor, but still significant (*45, 53, 54, 163, 176*). However, the role of A3F has been significant in some reports, but not others (*45, 50, 53, 98, 136, 164, 177*). In individual experiments with A3F, the mutational load is lower than A3G and A3H and this seemed incongruent with data demonstrating that A3F exerts a direct selective force on HIV-1 Vif and provides a protective effect in pretreatment HIV-1 cohorts (*4, 50, 51, 97, 164, 174*). Here we provide an explanation for these data demonstrating that the A3F 231V in comparison to the A3F 231I is more stable in cells, is partially resistant to Vif-mediated degradation, is more highly encapsidated into virions, induces more mutations, and can increase encapsidation of coexpressed A3G. These data suggest that the A3F 231V through its own action and hetero-oligomerization with A3F 231I or A3G is a contributor to G→A mutations in HIV-1+ individuals.

A protective effect of A3F 231V was demonstrated through genetic association analysis of pretreatment cohorts (*97*). Through this analysis it was shown that even when data are adjusted for known genetic factors (*HLA* allele and CCR5) the presence of even one A3F 231V allele can significantly delay progression to clinical AIDS and lower viral load (*97*). The authors suggested that this was due to partial resistance of A3F 231V to Vif, which had also been reported previously (*97, 136*). Here we show that the A3F 231V is protected from Vif mediated degradation in cells 7- 12- fold more than A3F 231I at high A3F transfection amounts (Figure 5.1E and Figure 5.2E) and A3F 231V is always detected in virions in the presence of Vif, even at low A3F transfection amounts where A3F 231I is not detected (Figure 5.2). However, the A3F 231V was not "resistant" to Vif mediated degradation, consistent with the 231 amino acid not being within the Vif binding interface (*167-170*). The protection from Vif appears to be due to the 5- to 10- fold higher steady state expression of A3F 231V in cells (-Vif, Figure 5.1C and Figure 5.2C). When equal amounts of A3F 231V and A3F 231I were expressed in cells, the capacity of Vif to induce degradation of both enzymes was similar (Figure 5.3). Rather, the data suggest that the higher steady state levels of A3F in cells exceeded the degradation capacity of Vif. Despite the protective effect of A3F

231V in HIV-1+ individuals, the partial protection of A3F 231V from Vif mediated degradation resulted in only a 20% decrease in infectivity in a Vif+ HIV-1 single-cycle infectivity experiment (Figure 5.4, Table 5.2). The A3F 231V did correlate with a 2- to 4- fold improved activity over A3F 231I to restrict ΔVif HIV-1 (Figure 5.1 and Figure 5.2). This was consistent with 2-fold more G→A mutations with A3F 231V than A3F 231I (Table 5.1). However, some previous studies have even not observed differences between the A3F 231V and A3F 231I variants (*89, 98*). Nonetheless, these small effects in cell-based studies still appear to be significant based on HIV-1+ pretreatment cohort (*97*). This situation is similar when considering the A3G H186R polymorphism. The A3G H186R polymorphism results in only a 2-fold decrease in A3G processivity and mutations induced in an *in vitro* assay (*58*), but is nonetheless linked with faster progression to AIDS in HIV-1+ individuals (*96, 166*).

A possible explanation for this is that the small effects observed in some cell-based studies may not be representative of the dynamics of multiple A3 enzymes in CD4+ T cells. We have previously shown that A3F 231I can hetero-oligomerize with A3G to produce an A3 complex with improved restriction activity against ΔVif HIV-1, consistent with other studies showing additive or synergistic effects of A3G and A3F (*33, 136, 178, 179*). However, this previously reported effect between A3G and A3F 231I had a different dynamic when A3G was coexpressed and hetero-oligomerized with A3F 231V (Figure 5.4, Figure 5.5, Table 5.2). Namely, the presence of A3F 231V with A3G increased A3G encapsidation in the presence of Vif (Figure 5.5). The coexpression and hetero-oligomerization of A3F 231V and A3F 231I increased the encapidation of A3F 231V 1.5-fold in the presence of Vif and resulted in a corresponding 1.7-fold increase in mutations and infectivity of 55% (Figure 5.5 and Table 5.3). Our data and previous analysis demonstrate that there are synergies between A3 enzymes that can enhance anti-HIV-1 activity, either directly due to changes in the biochemical characteristics of the enzymes or due to enhanced encapsidation in the presence of Vif (Figure 5.5) (*33*). Despite observations that A3F 231V on its own is partially protected from Vif-mediated degradation and our observation that A3F 231V appears to protect A3G from Vif-mediated degradation, the mechanism is not known (Figure 5.2 and Figure 5.5) (*97, 136*).

Our study suggests that the A3F 231V steady state expression levels exceed the Vif degradation capacity. Although Vif is expressed in sufficient time before virus assembly to purge cells of A3

enzymes, these biologically processes can never be considered complete as some A3 enzyme will always remain (*160*). Although the HIV-1 LTR is a strong promoter, having many transcription factor binding sites, the A3 enzymes are expressed constitutively in CD4+ T cells and expression is upregulated upon CD4+ T cell stimulation (*22, 48, 180*). In support of the idea that Vif-mediated degradation is incomplete, a study demonstrated that out of the functionally inactive HIV-1 proviral DNA in HIV-1+ individuals, A3-induced mutations accounted for inactivation of 20% of this pool, with the remaining amount being due to reverse transcriptase induced insertions and deletions (*87*). Since A3F requires ubiquitination on multiple sites for Vif-mediated degradation, it is unlikely that the A3F polymorphisms can alter ubiquitination and proteasome degradation kinetics (*181*). The A3F has also been shown to have equal or higher capacity to encapsidate into HIV-1 virions, especially within the core (*164, 182*). Thus, considering virion encapsidation one could conclude that different encapsidation mechanisms could be responsible, but in cells A3F 231V is protected from Vif-mediated degradation (Figure 5.1, Figure 5.2, and Figure 5.5). This together with the lower steady state levels of A3F 231I even when expressed from a plasmid suggest that there are structural stability differences between the A3F 231V and A3F 231I (Figure 5.1, Figure 5.2, and Figure 5.5). Altogether, the data support that the higher amount of A3F 231V in cells is not fully purged by HIV-1 Vif and warrants further study.

The reason for the stability differences between A3F 231V and A3F 231I remain to be elucidated. A nucleotide variant prediction model suggested that the A3F linked SNPs can have an effect on stability (*96*). The amino acid 108 is on predicted helix 3 in the noncatalytic N-terminal domain (NTD) and the amino acid 231 is on β-strand 2 in the catalytic C-terminal domain (*1, 168, 183, 184*). Neither of these structures are near the catalytic center, predicted nucleic acid binding motifs, or the Vif interface (*1, 167, 168, 183, 184*). Although V231I is a conservative amino acid change, a Val to Ile amino acid change in other proteins, e.g., transthyretin, has been found to destabilize equilibrium kinetics of oligomerization and result in physiological dysfunction (*185*). The A108S is a nonpolar to polar amino acid change and the Ser can form hydrogen bonds, unlike Ala, which may alter protein stability (*186*).

Numerous A3 polymporphisms have been identified that have affected HIV-1 restriction or restriction of retrotransposons (*89-91*). In regards to HIV-1 restriction, since lentiviral genomes have a mutation rate that far exceeds that of its host, a mechanism to combat this is by limiting the

genetic space for viral evolution through population and individual diversity in immune system polymorphisms (*187-189*). This has been demonstrated for A3G heterozygosity in Sabaeus African Green Monkeys where there exists forms of A3G that are sensitive or resistant to Vif-mediated degradation (*115*). The heterozygous animals could constrain Vif evolution since the evolution of Vif to induce degradation of one form of A3G, 130H, resulted in the concomitant sensitivity to the alternate allele, A3G 130D (*115*). Even in cell-based experiments, this same trade-off has been identified for HIV-1 engineered to be sensitive to A3s due to tandem stop codons in Vif (*174*). The evolution of this engineered HIV-1 to overcome A3G resulted in Vif independent mutations, but the acquired mutations to antagonize A3F were Vif specific and resulted in a sensitivity to A3G (*174, 190, 191*). As a result, the polymorphisms of A3s, within one paralog and between paralogs, can limit virus escape options. In the case of A3G and A3F, we have found that even the less active allele that codes for A3F 231I can promote A3G activity by increasing A3G processivity and delaying reverse transcription (*33*). This increased activity is important in the presence of Vif where A3G activity is decreased by 10-fold in a single-cycle infectivity assay (Table 5.1). More striking however was the maintenance of A3G encapsidation when coexpressed with A3F 231V and Vif+ HIV-1 at a level comparable to A3G alone with a ΔVif HIV-1 (Figure 5.5E). The disruption of Vif-mediated degradation through hetero-oligomerization of A3s appears to be an example of increasing host restriction factor diversity to combat virus evolution (*188, 189*). Consistent with genetic population analysis, the most noticeable example of this is when the A3F 231V and A3F 231I were coexpressed, as would occur in over 50% of the population (Figure 5.4, Figure 5.5, Table 5.2). In these individuals, that are protected against development of AIDS at an equivalent level to homozygous A3F 231V individuals (*97*), HIV-1 restriction in the presence of Vif was significant and resulted in a mutation rate only 2-fold less than A3F 231V expressed alone in a ΔVif HIV-1 single-cycle infectivity assay (Figure 5.5 and Table 5.3). Thus, although some polymorphisms may have small differences in the absence of Vif, their activity in the presence of Vif may be more significant.

Overall, our data demonstrate that the A3F 231V has higher steady state expression levels in cells and virion encapsidation than A3F 231I, in the presence and absence of Vif. This results in larger decreases in HIV-1 infectivity and higher mutation rates. The A3F 231V when coexpressed with A3G can enhance A3G encapsdiation and restriction activity. When A3F 231V is coexpressed with A3F 231I, which is a common heterozygosity in humans, HIV-1 restriction is enhanced in

the presence and absence of Vif. This supports a model in which not only the genetic diversity of the host for each A3, but the ability of these A3s to functionally interact can contribute to a diverse and multifaceted protective response during HIV-1 infection

6.0 General Discussion and Future Directions

6.1 General Discussion

The initial aims of this project began as a pilot study venturing outside familiar cell-based assays, and panning out to population-based experiments, working with primary cells and donor genotyping. Being able to start this project from its foundation allowed me to truly understand the realities and hardships that come along with optimizing and troubleshooting new experiments. The restriction digest genotyping assay is a prime example. From initial literature review, the assay described in Wang *et al.* 2011 was a simple and effective method for rapidly genotyping large numbers of donor SNPs in a cost-effective fashion (*90*). Although the methods were only described for A3H SNPs we used the same premise for assay design for A3G, A3F, and A3D SNPs. Due to a lack of reliable controls and in subjective error in interpreting results we chose to switch to Sanger Sequencing-based genotyping. The change in methods proved fruitful, with ease of data interpretation and reliable controls built into pre-sequence clone generation. The only reason we did not start with Sanger Sequencing was due to the projected volume of donor samples we expected to receive. However, with donor samples becoming limited this method became viable.

The genotyping results gave us insight into the A3F, A3G and A3H polymorphisms in our population, which in the APOBEC3 field had yet to be characterized in a mixed population setting. With allelic frequencies of A3F, A3G and A3H SNPs characterized in European Caucasian, East Asian and African lineages, general trends of SNP prevalence in each ethnicity can be used to theorize the SNP frequency in a mixed population using current demographics. Current demographics of Saskatchewan were drawn from Statistics Canada 2016 Census, informing us that our mixed population should consist of approximately 70.9% European Caucasian, 16.6 % Indigenous and 9.3% East Asian (Figure 2.1) (*124*). For A3H genotyping we saw allelic frequencies of active A3H forms (A3H hap II, hap V and hap VII) of 30%; predictive of a mixed population (Figure 3.3.4). The same trend of mixed-population predictive allelic frequencies can be said for both A3G and A3F, where we respectively see 8% and 50% of the less active variant (Figure 3.3.4, A3G R186, A3F 231I). When allelic frequencies are analyzed independently, our mixed population has a relatively low field of activity countering HIV-1 infection, with both high frequencies of inactive A3H alleles and less active A3F 231I alleles (Figure 3.3.4, Figure 5.4). Donors with similar A3H and A3F repertoires have been correlated to increased susceptibility to HIV-1 infection, lower $CD4^+$ T-cell counts, higher viral loads, and more rapid progression to AIDS through the course of infection (*81, 90, 92, 96*). However, when both alleles of each donor were

combined for whole genotype analysis we observed the genetic landscape of dominant A3F SNPs in our population drastically change. The dominant A3F genotype encoded for A3F 231V/231I heterozygotes (66%), compared to 231V/231V and 231I/231I homozygotes which both had a population prevalence of 17% (Figure 5.4B). Although it may seem counter intuitive, being heterozygous for alleles that vary in activity can increase the effectiveness of the restriction factor. This has been demonstrated for A3G heterozygosity in African Green Monkeys where there exists forms of A3G that are sensitive or resistant to Vif-mediated degradation; 130D and 130H respectively (*115*). The heterozygous animals resisted Vif evolution while homozygous conditions for both 130H and 130D alleles resulted in Vif sensitivity (*115*). This suggests that maintaining restriction factors with polymorphic regions can be functionally beneficial.

Our data supports restriction factor heterozygosity can functionally beneficial in human A3 repertoires A3G H186/H186 and A3F 231V/231I were dominant genotypes in our population, we then found that the A3F 231V variant when coexpressed with either A3G or A3F 231I resulted in an increase in cellular stability when expressed with both LAI and LAI Δvif HIV-1 (Figure 5.2). The rise in cellular expression resulted in a higher frequency of viral encapsidation when compared to A3G, A3F 231V and A3F 231I single expression systems alone (Figure 5.2). Importantly, under co-expression conditions A3F 231V was shown to hetero-oligomerize with both A3G and the A3F 231I SNP variant (Figure 5.4, Table 5.2).

The hetero-oligomerization of A3G/A3F 231V and A3G/A3F 231I resulted in the co-encapsidation of both enzymes allowing the enzymes to co-mutate the HIV-1 genome (Figure 5.2, Table 5.3). The rise in steady-state expression and encapsidation of A3F-231V when co expressed with A3G and A3F-231I also correlated to a rise in mutation frequency. That is, when A3F 231I-231V were co-expressed we observed a 3-fold increase in GA→AA mutations in comparison to A3F 231I alone (Table 5.3). When A3G was co-expressed with A3F 231V we observed a 1.4-fold increase in GG→AG mutation and a 2-fold increase in GA→AA mutations in comparison to A3G co-expressed with A3F 231I. These data indicated that A3F 231V when encapsidated with A3G or A3F 231I increased A3F mediated restriction through cooperative interaction (Table 5.3). Heterozygous A3F interactions also showed an increased resistance to HIV-1 Vif. When A3F 231V and A3F 231I were co-expressed in the presence of Vif we observed a 10-fold rise in GA→AA mutations in comparison to A3F 231I alone (Table 5.3).

The enhanced mutational capacity of coexpressed A3G and A3F 231V also translated into an increase in HIV-1 restriction. For A3F polymorphisms, the lowest percent infectivity was observed in the A3F 231V/A3F 231I co-expression condition of 18% at 50 ng (Figure 5.5) in the absence of Vif (Figure 5.5A). Mild Vif resistance was also observed from the A3F 231V-A3F 231I co-expression condition, where a 40% decline in infectivity was observed (Figure 5.5B).

A3F 231VI is a conservative amino acid change present on β-strand 2 in the catalytic C-terminal domain which does not localize to catalytic centers, predicted nucleic acid binding motifs, or the Vif interface (*1, 168, 183, 184*). Although, similar nucleotide variations in predictive models shows a link between stability and A3F SNPs. The reason for the stability infectivity differences between A3F 231V and A3F 231I remain to be elucidated (*96*). In the absence of Vif, A3G and A3G-A3F 231V conditions share the title of the most robust restrictors of HIV-1 replication, where we see nearly a 90% decline in infectivity at 50 ng for both conditions (Figure 5.5A).

6.2 Conclusions

From the data provided in our project, we see on a population level, individual allelic frequencies of A3 SNPs are more compromised than that of whole European Caucasian lineages, however novel APOBEC3 interactions elucidated by whole SNP genotypes are acting cooperatively to restrict retroviral restriction in the presence and absence of HIV-1 Vif. We confirm that A3F 231V is a more robust A3F variant than A3F 231I, that A3F 231V can modestly resist the function of HIV-1 Vif, when co-expressed with A3G and A3F 231I increases steady-state cellular expression, viral encapsidation and ssDNA mutagenesis; resulting in strong inter-A3 interactions that enhance retroviral restriction.

With HIV-1 infection rates being the highest in the country, investigation into route causes and treatment are of paramount importance. In order to properly treat this disease we must first understand it. This study shows that we have yet to completely understand the host-pathogen relationships held by humans and HIV-1, and with current trends of globalization, mixed population analyses have proven to elucidate novel interactions between our innate immune system and this rapidly revolving pathogen. Further examination of other genotype specific multi-A3 expression systems is the next logical step forward.

6.3 Future directions

From our donor PBMC genotyping and cell-based recreation of A3G and A3F repertoires, we have characterized novel A3 interactions between A3F polymorphisms and with A3G. In reality, A3D and A3H SNP variants are also expressed alongside the other A3 paralogs, thus a full repertoire of A3 SNP variants would include all cytoplasmic and retrovirally antagonistic A3 enzymes (A3F, A3G, A3H, and A3D). We have already characterized A3H genotypes in our population, and although A3D is a weak restrictor of HIV-1 replication we have the primers to amplify A3D polymorphic regions as well. Thus, the next logical step moving forward from this study would be to characterize A3D SNPs from our current population set, along with increasing our sample size for all A3 SNP genotypes and recreate a physiologically relevant transfection model to mimic HIV-1 replication restriction by a representative dominant population genotype in cells.

Once we have established prevalence of all A3 SNPs we could then begin to screen donors for specific combinations of A3 SNPs and monitor HIV-1 replication kinetics using Vif+ HIV-1 strains in human PBMCs. Deaminations and mutations could be monitored through time and full A3 repertoires could be analyzed for replication restriction efficiency under physiologically relevant conditions.

In our previous study we also show that by titrating in greater amounts of A3 coding plasmid in transfections, we see a negative correlation with viral infectivity. That is as A3 expression titer increase, viral infectivity decreases, that is, if more A3 is expressed in cells more is encapsidated; which increases the frequency of mutation thereby decrease viral infectivity. Thus, we are also interested in doing qPCR analysis of A3 transcripts between donors to determine if there is any variation in steady state A3 expression between donors. Finding variations between donors could indicate similar results to increasing A3 transfection amounts, where we may see enhanced retroviral restriction ability from different donors.

References

1. Y. Feng, T. T. Baig, R. P. Love, L. Chelico, Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. *Front Microbiol* **5**, 450 (2014).
2. Y. L. Chiu, W. C. Greene, The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol* **26**, 317-353 (2008).
3. R. S. Harris *et al.*, DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803-809 (2003).
4. M. B. Adolph, R. P. Love, L. Chelico, Biochemical Basis of APOBEC3 Deoxycytidine Deaminase Activity on Diverse DNA Substrates. *ACS Infect Dis* **4**, 224-238 (2018).
5. M. Muramatsu *et al.*, Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *The Journal of biological chemistry* **274**, 18470-18476 (1999).
6. R. S. LaRue *et al.*, The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC Mol Biol* **9**, 104 (2008).
7. H. Ohtsubo *et al.*, APOBEC2 negatively regulates myoblast differentiation in muscle regeneration. *Int J Biochem Cell Biol* **85**, 91-101 (2017).
8. A. Jarmuz *et al.*, An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79**, 285-296 (2002).
9. B. R. Cullen, Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. *J Virol* **80**, 1067-1076 (2006).
10. M. Lochelt *et al.*, The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. *Proc Natl Acad Sci U S A* **102**, 7982-7987 (2005).
11. J. A. Hayward *et al.*, Differential Evolution of Antiretroviral Restriction Factors in Pteropid Bats as Revealed by APOBEC3 Gene Complexity. *Mol Biol Evol* **35**, 1626-1637 (2018).
12. J. E. Wedekind, G. S. Dance, M. P. Sowden, H. C. Smith, Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet* **19**, 207-216 (2003).
13. S. L. Sawyer, M. Emerman, H. S. Malik, Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol* **2**, E275 (2004).
14. L. Etienne *et al.*, The Role of the Antiviral APOBEC3 Gene Family in Protecting Chimpanzees against Lentiviruses from Monkeys. *PLoS pathogens* **11**, e1005149 (2015).
15. Y. Nakano *et al.*, A conflict of interest: the evolutionary arms race between mammalian APOBEC3 and lentiviral Vif. *Retrovirology* **14**, 31 (2017).
16. R. S. Harris, J. P. Dudley, APOBECs and virus restriction. *Virology* **479-480**, 131-145 (2015).
17. J. D. Salter, R. P. Bennett, H. C. Smith, The APOBEC Protein Family: United by Structure, Divergent in Function. *Trends Biochem Sci* **41**, 578-594 (2016).
18. M. L. Hegde, T. K. Hazra, S. Mitra, Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res* **18**, 27-47 (2008).
19. Y. J. Kim, D. M. Wilson, 3rd, Overview of base excision repair biochemistry. *Curr Mol Pharmacol* **5**, 3-13 (2012).

20. A. B. Robertson, A. Klungland, T. Rognes, I. Leiros, DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci* **66**, 981-993 (2009).
21. B. Demple, J. S. Sung, Molecular and biological roles of Ape1 protein in mammalian base excision repair. *DNA Repair (Amst)* **4**, 1442-1449 (2005).
22. E. W. Refsland *et al.*, Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res* **38**, 4274-4284 (2010).
23. H. E. Krokan, F. Drablos, G. Slupphaug, Uracil in DNA--occurrence, consequences and repair. *Oncogene* **21**, 8935-8948 (2002).
24. A. Koito, T. Ikeda, Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front Microbiol* **4**, 28 (2013).
25. K. Belanger, M. Savoie, M. C. Rosales Gerpe, J. F. Couture, M. A. Langlois, Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. *Nucleic Acids Res* **41**, 7438-7452 (2013).
26. Y. Iwatani *et al.*, Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res* **35**, 7096-7108 (2007).
27. V. Palchevskiy, S. E. Finkel, A role for single-stranded exonucleases in the use of DNA as a nutrient. *J Bacteriol* **191**, 3712-3716 (2009).
28. Y. Feng, M. H. Goubran, T. B. Follack, L. Chelico, Deamination-independent restriction of LINE-1 retrotransposition by APOBEC3H. *Sci Rep* **7**, 10881 (2017).
29. T. Kobayashi *et al.*, Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. *Journal of virology* **88**, 5881-5887 (2014).
30. K. N. Bishop, R. K. Holmes, M. H. Malim, Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. *Journal of virology* **80**, 8450-8458 (2006).
31. K. N. Bishop, M. Verma, E. Y. Kim, S. M. Wolinsky, M. H. Malim, APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS pathogens* **4**, e1000231 (2008).
32. R. K. Holmes, F. A. Koning, K. N. Bishop, M. H. Malim, 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 (2007).
33. A. Ara *et al.*, Mechanism of Enhanced HIV Restriction by Virion Coencapsidated Cytidine Deaminases APOBEC3F and APOBEC3G. *J Virol* **91**, (2017).
34. D. Pollpeter *et al.*, Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G. *Nat Microbiol* **3**, 220-233 (2018).
35. J. I. Hoopes *et al.*, APOBEC3A and APOBEC3B Preferentially Deaminate the Lagging Strand Template during DNA Replication. *Cell Rep* **14**, 1273-1282 (2016).
36. V. B. Seplyarskiy *et al.*, APOBEC-induced mutations in human cancers are strongly enriched on the lagging DNA strand during replication. *Genome Res* **26**, 174-182 (2016).
37. K. Collins, T. Jacks, N. P. Pavletich, The cell cycle and cancer. *Proc Natl Acad Sci U S A* **94**, 2776-2778 (1997).
38. T. Sandal, Molecular aspects of the mammalian cell cycle and cancer. *Oncologist* **7**, 73-81 (2002).
39. N. Kanu *et al.*, DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. *Genome Biol* **17**, 185 (2016).
40. S. A. Roberts *et al.*, An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet* **45**, 970-976 (2013).

41. C. Swanton, N. McGranahan, G. J. Starrett, R. S. Harris, APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. *Cancer Discov* **5**, 704-712 (2015).
42. H. C. Smith, R. P. Bennett, A. Kizilyer, W. M. McDougall, K. M. Prohaska, Functions and regulation of the APOBEC family of proteins. *Seminars in cell & developmental biology* **23**, 258-268 (2012).
43. G. J. Starrett *et al.*, The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *Nat Commun* **7**, 12918 (2016).
44. M. B. Burns *et al.*, APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* **494**, 366-370 (2013).
45. J. F. Hultquist *et al.*, Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J Virol* **85**, 11220-11234 (2011).
46. Y. Dang, X. Wang, W. J. Esselman, Y. H. Zheng, Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *Journal of virology* **80**, 10522-10533 (2006).
47. L. Lackey, E. K. Law, W. L. Brown, R. S. Harris, Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell cycle (Georgetown, Tex.)* **12**, 762-772 (2013).
48. F. A. Koning *et al.*, Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *Journal of virology* **83**, 9474-9485 (2009).
49. M. B. Adolph, J. Webb, L. Chelico, Retroviral restriction factor APOBEC3G delays the initiation of DNA synthesis by HIV-1 reverse transcriptase. *PLoS One* **8**, e64196 (2013).
50. A. Ara, R. P. Love, L. Chelico, Different mutagenic potential of HIV-1 restriction factors APOBEC3G and APOBEC3F is determined by distinct single-stranded DNA scanning mechanisms. *PLoS Pathog* **10**, e1004024 (2014).
51. Y. Feng *et al.*, Natural Polymorphisms and Oligomerization of Human APOBEC3H Contribute to Single-stranded DNA Scanning Ability. *J Biol Chem* **290**, 27188-27203 (2015).
52. M. B. Adolph *et al.*, Cytidine deaminase efficiency of the lentiviral viral restriction factor APOBEC3C correlates with dimerization. *Nucleic Acids Res* **45**, 3378-3394 (2017).
53. C. Chaipan, J. L. Smith, W. S. Hu, V. K. Pathak, APOBEC3G restricts HIV-1 to a greater extent than APOBEC3F and APOBEC3DE in human primary CD4+ T cells and macrophages. *Journal of virology* **87**, 444-453 (2013).
54. E. W. Refsland, J. F. Hultquist, R. S. Harris, Endogenous origins of HIV-1 G-to-A hypermutation and restriction in the nonpermissive T cell line CEM2n. *PLoS Pathog* **8**, e1002800 (2012).
55. H. Huthoff, F. Autore, S. Gallois-Montbrun, F. Fraternali, M. H. Malim, RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1. *PLoS pathogens* **5**, e1000330 (2009).
56. L. Chelico, E. J. Sacho, D. A. Erie, M. F. Goodman, A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. *J Biol Chem* **283**, 13780-13791 (2008).
57. L. Apolonia *et al.*, Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. *PLoS pathogens* **11**, e1004609 (2015).

58. Y. Feng, L. Chelico, Intensity of deoxycytidine deamination of HIV-1 proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the noncatalytic domain. *J Biol Chem* **286**, 11415-11426 (2011).
59. J. Fettig, M. Swaminathan, C. S. Murrill, J. E. Kaplan, Global epidemiology of HIV. *Infect Dis Clin North Am* **28**, 323-337 (2014).
60. D. D. Richman, Introduction: challenges to finding a cure for HIV infection. *Curr Opin HIV AIDS* **6**, 1-3 (2011).
61. B. G. Turner, M. F. Summers, Structural biology of HIV. *J Mol Biol* **285**, 1-32 (1999).
62. D. A. Jacques *et al.*, HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349-353 (2016).
63. J. G. Levin, M. Mitra, A. Mascarenhas, K. Musier-Forsyth, Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biol* **7**, 754-774 (2010).
64. A. York, S. B. Kutluay, M. Errando, P. D. Bieniasz, The RNA Binding Specificity of Human APOBEC3 Proteins Resembles That of HIV-1 Nucleocapsid. *PLoS pathogens* **12**, e1005833 (2016).
65. W. E. Johnson, R. C. Desrosiers, Viral persistence: HIV's strategies of immune system evasion. *Annu Rev Med* **53**, 499-518 (2002).
66. D. S. Burke, Recombination in HIV: an important viral evolutionary strategy. *Emerg Infect Dis* **3**, 253-259 (1997).
67. J. D. Roberts, K. Bebenek, T. A. Kunkel, The accuracy of reverse transcriptase from HIV-1. *Science (New York, N.Y.)* **242**, 1171-1173 (1988).
68. K. Bebenek, J. Abbotts, S. H. Wilson, T. A. Kunkel, Error-prone polymerization by HIV-1 reverse transcriptase. Contribution of template-primer misalignment, miscoding, and termination probability to mutational hot spots. *The Journal of biological chemistry* **268**, 10324-10334 (1993).
69. E. Y. Kim *et al.*, Human APOBEC3 induced mutation of human immunodeficiency virus type-1 contributes to adaptation and evolution in natural infection. *PLoS pathogens* **10**, e1004281 (2014).
70. B. A. Desimmie *et al.*, Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. *J Mol Biol* **426**, 1220-1245 (2014).
71. B. J. Stanley *et al.*, Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly. *Journal of virology* **82**, 8656-8663 (2008).
72. Y. Yu, Z. Xiao, E. S. Ehrlich, X. Yu, X. F. Yu, Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. *Genes Dev* **18**, 2867-2872 (2004).
73. Z. Xiao *et al.*, Assembly of HIV-1 Vif-Cul5 E3 ubiquitin ligase through a novel zinc-binding domain-stabilized hydrophobic interface in Vif. *Virology* **349**, 290-299 (2006).
74. X. Yu *et al.*, Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056-1060 (2003).
75. W. Zhang, J. Du, S. L. Evans, Y. Yu, X. F. Yu, T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* **481**, 376-379 (2011).
76. R. Yoshikawa *et al.*, Vif determines the requirement for CBF-beta in APOBEC3 degradation. *J Gen Virol* **96**, 887-892 (2015).
77. S. Jager *et al.*, Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. *Nature* **481**, 371-375 (2011).

78. X. Han *et al.*, Evolutionarily conserved requirement for core binding factor beta in the assembly of the human immunodeficiency virus/simian immunodeficiency virus Vif-cullin 5-RING E3 ubiquitin ligase. *Journal of virology* **88**, 3320-3328 (2014).
79. D. Wei, Y. Sun, Small RING Finger Proteins RBX1 and RBX2 of SCF E3 Ubiquitin Ligases: The Role in Cancer and as Cancer Targets. *Genes Cancer* **1**, 700-707 (2010).
80. M. Ooms, M. Letko, M. Binka, V. Simon, The resistance of human APOBEC3H to HIV-1 NL4-3 molecular clone is determined by a single amino acid in Vif. *PLoS One* **8**, e57744 (2013).
81. E. W. Refsland *et al.*, Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity. *PLoS Genet* **10**, e1004761 (2014).
82. M. Ooms *et al.*, HIV-1 Vif adaptation to human APOBEC3H haplotypes. *Cell Host Microbe* **14**, 411-421 (2013).
83. M. Binka, M. Ooms, M. Steward, V. Simon, The activity spectrum of Vif from multiple HIV-1 subtypes against APOBEC3G, APOBEC3F, and APOBEC3H. *Journal of virology* **86**, 49-59 (2012).
84. A. Mehle *et al.*, Identification of an APOBEC3G binding site in human immunodeficiency virus type 1 Vif and inhibitors of Vif-APOBEC3G binding. *Journal of virology* **81**, 13235-13241 (2007).
85. R. A. Russell, V. K. Pathak, Identification of two distinct human immunodeficiency virus type 1 Vif determinants critical for interactions with human APOBEC3G and APOBEC3F. *Journal of virology* **81**, 8201-8210 (2007).
86. B. Schrofelbauer, T. Senger, G. Manning, N. R. Landau, Mutational alteration of human immunodeficiency virus type 1 Vif allows for functional interaction with nonhuman primate APOBEC3G. *Journal of virology* **80**, 5984-5991 (2006).
87. K. M. Bruner *et al.*, Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat Med* **22**, 1043-1049 (2016).
88. V. Simon *et al.*, Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS Pathog* **1**, e6 (2005).
89. N. K. Duggal, W. Fu, J. M. Akey, M. Emerman, Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations. *Virology* **443**, 329-337 (2013).
90. X. Wang *et al.*, Analysis of human APOBEC3H haplotypes and anti-human immunodeficiency virus type 1 activity. *J Virol* **85**, 3142-3152 (2011).
91. M. OhAinle, J. A. Kerns, M. M. Li, H. S. Malik, M. Emerman, Antiretroelement activity of APOBEC3H was lost twice in recent human evolution. *Cell Host Microbe* **4**, 249-259 (2008).
92. D. Sakurai *et al.*, APOBEC3H polymorphisms associated with the susceptibility to HIV-1 infection and AIDS progression in Japanese. *Immunogenetics* **67**, 253-257 (2015).
93. H. R. Rangel *et al.*, Deletion, insertion and stop codon mutations in vif genes of HIV-1 infecting slow progressor patients. *J Infect Dev Ctries* **3**, 531-538 (2009).
94. P. Sova, D. J. Volsky, L. Wang, W. Chao, Vif is largely absent from human immunodeficiency virus type 1 mature virions and associates mainly with viral particles containing unprocessed gag. *Journal of virology* **75**, 5504-5517 (2001).

95. K. Reddy *et al.*, APOBEC3G expression is dysregulated in primary HIV-1 infection and polymorphic variants influence CD4+ T-cell counts and plasma viral load. *AIDS* **24**, 195-204 (2010).
96. P. An *et al.*, APOBEC3G genetic variants and their influence on the progression to AIDS. *J Virol* **78**, 11070-11076 (2004).
97. P. An *et al.*, Role of APOBEC3F Gene Variation in HIV-1 Disease Progression and Pneumocystis Pneumonia. *PLoS Genet* **12**, e1005921 (2016).
98. L. C. Mulder *et al.*, Moderate influence of human APOBEC3F on HIV-1 replication in primary lymphocytes. *Journal of virology* **84**, 9613-9617 (2010).
99. J. Gu *et al.*, Biochemical Characterization of APOBEC3H Variants: Implications for Their HIV-1 Restriction Activity and mC Modification. *J Mol Biol* **428**, 4626-4638 (2016).
100. S. Lopez, L. van Dorp, G. Hellenthal, Human Dispersal Out of Africa: A Lasting Debate. *Evol Bioinform Online* **11**, 57-68 (2015).
101. B. P. McEvoy, J. E. Powell, M. E. Goddard, P. M. Visscher, Human population dispersal "Out of Africa" estimated from linkage disequilibrium and allele frequencies of SNPs. *Genome Res* **21**, 821-829 (2011).
102. M. A. Papathanasopoulos, G. M. Hunt, C. T. Tiemessen, Evolution and diversity of HIV-1 in Africa--a review. *Virus genes* **26**, 151-163 (2003).
103. P. M. Sharp, B. H. Hahn, The evolution of HIV-1 and the origin of AIDS. *Philos Trans R Soc Lond B Biol Sci* **365**, 2487-2494 (2010).
104. B. H. Hahn, G. M. Shaw, K. M. De Cock, P. M. Sharp, AIDS as a zoonosis: scientific and public health implications. *Science (New York, N.Y.)* **287**, 607-614 (2000).
105. L. Etienne, B. H. Hahn, P. M. Sharp, F. A. Matsen, M. Emerman, Gene loss and adaptation to hominids underlie the ancient origin of HIV-1. *Cell Host Microbe* **14**, 85-92 (2013).
106. N. K. Duggal, M. Emerman, Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol* **12**, 687-695 (2012).
107. W. E. Johnson, S. L. Sawyer, Molecular evolution of the antiretroviral TRIM5 gene. *Immunogenetics* **61**, 163-176 (2009).
108. J. S. Albin, R. S. Harris, Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: implications for therapeutics. *Expert Rev Mol Med* **12**, e4 (2010).
109. H. Lahouassa *et al.*, SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol* **13**, 223-228 (2012).
110. D. Wolf, S. P. Goff, Host restriction factors blocking retroviral replication. *Annu Rev Genet* **42**, 143-163 (2008).
111. E. Bailes *et al.*, Hybrid origin of SIV in chimpanzees. *Science (New York, N.Y.)* **300**, 1713 (2003).
112. J. L. Heeney, A. G. Dalgleish, R. A. Weiss, Origins of HIV and the evolution of resistance to AIDS. *Science (New York, N.Y.)* **313**, 462-466 (2006).
113. N. Wood *et al.*, HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS pathogens* **5**, e1000414 (2009).
114. J. C. Venter, H. O. Smith, M. D. Adams, The Sequence of the Human Genome. *Clin Chem* **61**, 1207-1208 (2015).

115. A. A. Compton, V. M. Hirsch, M. Emerman, The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. *Cell Host Microbe* **11**, 91-98 (2012).
116. C. Tian *et al.*, Differential requirement for conserved tryptophans in human immunodeficiency virus type 1 Vif for the selective suppression of APOBEC3G and APOBEC3F. *Journal of virology* **80**, 3112-3115 (2006).
117. J. S. Albin, W. L. Brown, R. S. Harris, Catalytic activity of APOBEC3F is required for efficient restriction of Vif-deficient human immunodeficiency virus. *Virology* **450-451**, 49-54 (2014).
118. J. L. Anderson, T. J. Hope, APOBEC3G restricts early HIV-1 replication in the cytoplasm of target cells. *Virology* **375**, 1-12 (2008).
119. C. M. Horvath, J. E. Darnell, Jr., The antiviral state induced by alpha interferon and gamma interferon requires transcriptionally active Stat1 protein. *Journal of virology* **70**, 647-650 (1996).
120. D. E. Levy, A. Garcia-Sastre, The virus battles: IFN induction of the antiviral state and mechanisms of viral evasion. *Cytokine Growth Factor Rev* **12**, 143-156 (2001).
121. S. G. Conticello, The AID/APOBEC family of nucleic acid mutators. *Genome Biol* **9**, 229 (2008).
122. R. Nowarski, M. Kotler, APOBEC3 cytidine deaminases in double-strand DNA break repair and cancer promotion. *Cancer Res* **73**, 3494-3498 (2013).
123. E. M. Bourgeois AC, Awan A, Jonah L, Varsaneux O, Siu W, HIV in Canada: 2016. *Public Health Agency of Canada (PHAC) - Canada Communicable Disease Reports* **43**, 248-256 (2017).
124. S. Canada, Focus on Geography Series. *2016 Census* **98-404-X2016001**, (2017).
125. P. Artimo *et al.*, ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* **40**, W597-603 (2012).
126. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
127. P. P. Rose, B. T. Korber, Detecting hypermutations in viral sequences with an emphasis on G --> A hypermutation. *Bioinformatics* **16**, 400-401 (2000).
128. S. Grossmann, P. Nowak, U. Neogi, Subtype-independent near full-length HIV-1 genome sequencing and assembly to be used in large molecular epidemiological studies and clinical management. *J Int AIDS Soc* **18**, 20035 (2015).
129. Z. L. Brumme *et al.*, Reduced replication capacity of NL4-3 recombinant viruses encoding reverse transcriptase-integrase sequences from HIV-1 elite controllers. *Journal of acquired immune deficiency syndromes (1999)* **56**, 100-108 (2011).
130. T. Vincze, J. Posfai, R. J. Roberts, NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res* **31**, 3688-3691 (2003).
131. R. S. Harris, J. F. Hultquist, D. T. Evans, The restriction factors of human immunodeficiency virus. *The Journal of biological chemistry* **287**, 40875-40883 (2012).
132. B. Mangeat *et al.*, Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**, 99-103 (2003).
133. H. Zhang *et al.*, The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94-98 (2003).
134. R. S. Harris, S. K. Petersen-Mahrt, M. S. Neuberger, RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* **10**, 1247-1253 (2002).

135. A. M. Sheehy, N. C. Gaddis, J. D. Choi, M. H. Malim, Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646-650 (2002).
136. M. T. Liddament, W. L. Brown, A. J. Schumacher, R. S. Harris, APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. *Curr Biol* **14**, 1385-1391 (2004).
137. H. L. Wiegand, B. P. Doehle, H. P. Bogerd, B. R. Cullen, A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J* **23**, 2451-2458 (2004).
138. Y. H. Zheng *et al.*, Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *Journal of virology* **78**, 6073-6076 (2004).
139. C. J. Wittkopp, M. B. Adolph, L. I. Wu, L. Chelico, M. Emerman, A Single Nucleotide Polymorphism in Human APOBEC3C Enhances Restriction of Lentiviruses. *PLoS pathogens* **12**, e1005865 (2016).
140. Q. Yu *et al.*, Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat Struct Mol Biol* **11**, 435-442 (2004).
141. S. G. Conticello, R. S. Harris, M. S. Neuberger, The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr Biol* **13**, 2009-2013 (2003).
142. S. Kao *et al.*, The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity. *Journal of virology* **77**, 11398-11407 (2003).
143. R. Mariani *et al.*, Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* **114**, 21-31 (2003).
144. M. Marin, K. M. Rose, S. L. Kozak, D. Kabat, HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat Med* **9**, 1398-1403 (2003).
145. A. M. Sheehy, N. C. Gaddis, M. H. Malim, The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature medicine* **9**, 1404-1407 (2003).
146. K. Stopak, C. de Noronha, W. Yonemoto, W. C. Greene, HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Molecular cell* **12**, 591-601 (2003).
147. Y. Guo *et al.*, Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* **505**, 229-233 (2014).
148. S. Jager *et al.*, Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. *Nature* **481**, 371-375 (2012).
149. M. Janini, M. Rogers, D. R. Birx, F. E. McCutchan, Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. *J Virol* **75**, 7973-7986 (2001).
150. J. P. Vartanian, M. Henry, S. Wain-Hobson, Sustained G-->A hypermutation during reverse transcription of an entire human immunodeficiency virus type 1 strain Vau group O genome. *J Gen Virol* **83**, 801-805 (2002).
151. J. P. Vartanian, A. Meyerhans, B. Asjö, S. Wain-Hobson, Selection, recombination, and G----A hypermutation of human immunodeficiency virus type 1 genomes. *J Virol* **65**, 1779-1788 (1991).

152. J. P. Vartanian, A. Meyerhans, M. Sala, S. Wain-Hobson, G-->A hypermutation of the human immunodeficiency virus type 1 genome: evidence for dCTP pool imbalance during reverse transcription. *Proc Natl Acad Sci U S A* **91**, 3092-3096 (1994).
153. V. K. Pathak, H. M. Temin, Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. *Proc Natl Acad Sci U S A* **87**, 6019-6023 (1990).
154. T. L. Kieffer *et al.*, G-->A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *J Virol* **79**, 1975-1980 (2005).
155. C. Pace *et al.*, Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. *Journal of virology* **80**, 9259-9269 (2006).
156. A. Piantadosi, D. Humes, B. Chohan, R. S. McClelland, J. Overbaugh, Analysis of the percentage of human immunodeficiency virus type 1 sequences that are hypermutated and markers of disease progression in a longitudinal cohort, including one individual with a partially defective Vif. *J Virol* **83**, 7805-7814 (2009).
157. M. L. de Lima-Stein *et al.*, In vivo HIV-1 hypermutation and viral loads among antiretroviral-naive Brazilian patients. *AIDS Res Hum Retroviruses* **30**, 867-880 (2014).
158. L. M. Eyzaguirre *et al.*, Elevated hypermutation levels in HIV-1 natural viral suppressors. *Virology* **443**, 306-312 (2013).
159. A. M. Land *et al.*, Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. *J Virol* **82**, 8172-8182 (2008).
160. M. Holmes, F. Zhang, P. D. Bieniasz, Single-Cell and Single-Cycle Analysis of HIV-1 Replication. *PLoS Pathog* **11**, e1004961 (2015).
161. M. A. Langlois, R. C. Beale, S. G. Conticello, M. S. Neuberger, 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 (2005).
162. T. T. Baig, Y. Feng, L. Chelico, Determinants of efficient degradation of APOBEC3 restriction factors by HIV-1 Vif. *J Virol* **88**, 14380-14395 (2014).
163. N. K. Duggal, H. S. Malik, M. Emerman, The breadth of antiviral activity of Apobec3DE in chimpanzees has been driven by positive selection. *J Virol* **85**, 11361-11371 (2011).
164. V. Zennou, P. D. Bieniasz, Comparative analysis of the antiretroviral activity of APOBEC3G and APOBEC3F from primates. *Virology* **349**, 31-40 (2006).
165. M. Ooms, S. Majdak, C. W. Seibert, A. Harari, V. Simon, The localization of APOBEC3H variants in HIV-1 virions determines their antiviral activity. *Journal of virology* **84**, 7961-7969 (2010).
166. K. Reddy *et al.*, Functional characterization of Vif proteins from HIV-1 infected patients with different APOBEC3G haplotypes. *AIDS* **30**, 1723-1729 (2016).
167. C. Richards *et al.*, The Binding Interface between Human APOBEC3F and HIV-1 Vif Elucidated by Genetic and Computational Approaches. *Cell Rep* **13**, 1781-1788 (2015).
168. M. Nakashima *et al.*, Structural Insights into HIV-1 Vif-APOBEC3F Interaction. *Journal of virology* **90**, 1034-1047 (2016).

169. J. S. Albin *et al.*, A single amino acid in human APOBEC3F alters susceptibility to HIV-1 Vif. *The Journal of biological chemistry* **285**, 40785-40792 (2010).
170. J. L. Smith, V. K. Pathak, Identification of specific determinants of human APOBEC3F, APOBEC3C, and APOBEC3DE and African green monkey APOBEC3F that interact with HIV-1 Vif. *Journal of virology* **84**, 12599-12608 (2010).
171. Y. Feng, R. P. Love, L. Chelico, HIV-1 viral infectivity factor (Vif) alters processive single-stranded DNA scanning of the retroviral restriction factor APOBEC3G. *J Biol Chem* **288**, 6083-6094 (2013).
172. E. Britan-Rosich, R. Nowarski, M. Kotler, Multifaceted counter-APOBEC3G mechanisms employed by HIV-1 Vif. *J Mol Biol* **410**, 1065-1076 (2011).
173. G. Hache, M. T. Liddament, R. S. Harris, The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. *J Biol Chem* **280**, 10920-10924 (2005).
174. J. S. Albin, G. Hache, J. F. Hultquist, W. L. Brown, R. S. Harris, Long-term restriction by APOBEC3F selects human immunodeficiency virus type 1 variants with restored Vif function. *Journal of virology* **84**, 10209-10219 (2010).
175. J. L. Mbisa, W. Bu, V. K. Pathak, APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. *J Virol* **84**, 5250-5259 (2010).
176. Y. Dang *et al.*, Identification of a single amino acid required for APOBEC3 antiretroviral cytidine deaminase activity. *Journal of virology* **85**, 5691-5695 (2011).
177. E. Miyagi *et al.*, Stably expressed APOBEC3F has negligible antiviral activity. *Journal of virology* **84**, 11067-11075 (2010).
178. B. A. Desimmie *et al.*, APOBEC3 proteins can copackage and comutate HIV-1 genomes. *Nucleic Acids Res* **44**, 7848-7865 (2016).
179. J. F. Krisko, F. Martinez-Torres, J. L. Foster, J. V. Garcia, HIV restriction by APOBEC3 in humanized mice. *PLoS pathogens* **9**, e1003242 (2013).
180. J. Karn, C. M. Stoltzfus, Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med* **2**, a006916 (2012).
181. J. S. Albin *et al.*, Dispersed sites of HIV Vif-dependent polyubiquitination in the DNA deaminase APOBEC3F. *J Mol Biol* **425**, 1172-1182 (2013).
182. C. Song, L. Sutton, M. E. Johnson, R. T. D'Aquila, J. P. Donahue, Signals in APOBEC3F N-terminal and C-terminal deaminase domains each contribute to encapsidation in HIV-1 virions and are both required for HIV-1 restriction. *The Journal of biological chemistry* **287**, 16965-16974 (2012).
183. K. K. Siu, A. Sultana, F. C. Azimi, J. E. Lee, Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. *Nat Commun* **4**, 2593 (2013).
184. M. F. Bohn *et al.*, Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. *Structure (London, England : 1993)* **21**, 1042-1050 (2013).
185. X. Jiang, J. N. Buxbaum, J. W. Kelly, The V122I cardiomyopathy variant of transthyretin increases the velocity of rate-limiting tetramer dissociation, resulting in accelerated amyloidosis. *Proceedings of the National Academy of Sciences* **98**, 14943-14948 (2001).
186. D. Nelson, Lehninger, Albert L, & Cox, Michael M., *Lehninger principles of biochemistry*. (W.H. Freeman., New York, ed. 5th ed, 2008).
187. M. Carrington *et al.*, HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* **283**, 1748-1752 (1999).

188. N. R. Meyerson, S. L. Sawyer, Two-stepping through time: mammals and viruses. *Trends Microbiol* **19**, 286-294 (2011).
189. A. A. Compton, H. S. Malik, M. Emerman, Host gene evolution traces the evolutionary history of ancient primate lentiviruses. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120496 (2013).
190. G. Hache, K. Shindo, J. S. Albin, R. S. Harris, Evolution of HIV-1 isolates that use a novel Vif-independent mechanism to resist restriction by human APOBEC3G. *Curr Biol* **18**, 819-824 (2008).
191. G. Hache, T. E. Abbink, B. Berkhout, R. S. Harris, Optimal translation initiation enables Vif-deficient human immunodeficiency virus type 1 to escape restriction by APOBEC3G. *Journal of virology* **83**, 5956-5960 (2009).

Appendix

Table A.1

Primer sequences for A3G, A3F, A3D and A3H amplification for use in restriction digest genotyping assay. A3H primers drawn from Wang *et al.* 2011 (*90*).

SNP amplicon	Primer Sequence
A3G-H186R F	AATTTCAGCACTGTTGGAGC
A3G-H186R R	GAGAATCTCCCCCAGCAT
A3F A108S F	GTGTATTCCCAGCCTGAGCA
A3F A108S R	CTTCATCGTCCATAATCTTCACG
A3F 231IV F	AGCCTATGGTCGGAACGAAA
A3F 231IV R	CTGGTTTCGGAAGACGCC
A3D R97C F	GTGTATTTCCGGTTTGAGAACC
A3D R97C R	GATGGTCAGGGTGACATTGG
A3D R248K F	ACCCGATGGAGGCAA
A3D R248K R	CTGGTTTCGGAAGAC
A3H N15Δ F	CCGAAACATTCCGCTTACAG
A3H N15Δ R	GCGTCAGCTGGTAACACAAG
A3H G105R F	CATGGGACTGGACGAAGCGCA
A3H G105R R	TGGGATCCACACAGAAGCCGCA
A3H 140K F	GCATCTTCGCCTCCCGCCTGT
A3H 140K R	CCCTGCCAAGCCTGTGCCC
A3H R18L F	TGGGTTTGAAAAGTGGCTTG
A3H R18L R	AACTGGGCCACTCAGATCC
A3H K121D F	CCACGCACTAGAAAGTTCACC

A3H K121D R	GGTTTGCACTCTTATAACTGCAAAG
A3H E178D F	CAAGAGCTCCTGGCACTG
A3H E178D R	CTGGTGCGAGAGGAGAACAC

Table A.2

Templates for A3G, A3F, A3D, and A3H restriction digest genotyping assay. Indicating each SNP, restriction enzyme and proposed uncut and cut banding patterns for both homozygous and heterozygous conditions. All values are in base pairs (bp)

Table A.2A

APOBEC3H N15Δ	15N/15N	15N/15Δ	15Δ/15Δ	Tm: 60
	90	90		
		87	87	

Table A.2B

APOBEC3H R18L	Uncut	18L/18L	18L/18R	18R/18R	Tm: 61
w/Fnu4HI	418				
Control:		222	222	222	
Uncut = 40		196	196		
Cut = 20			141	141	
			55	55	

Table A.2C

APOBEC3H G105R	Uncut	105G/105G	105G/105R	105R/105R	Tm: 72
w/ HhaI	197				
Control:		180	180		
Uncut = 40			102	102	
Cut = 20			78	78	
		17	17	17	

Table A.2D

APOBEC3H K121D	Uncut	121D/121D	121D/121K	121K/121K	Tm: 60
w/HpyAV	441				
Control:		322	322		
Uncut = 40			220	220	
Cut = 20		119	119	119	
			102	102	

Table A.2E

APOBEC3H E178D	Uncut	178D/178D	178D/178E	178E/178E	Tm: 60
w/SmlI	636				
Control:		530	530		
Uncut = 40		353	353		
Cut = 20			177	177	
		106	106	106	

Table A.2F

APOBEC3G H186R	Uncut	186H/186H	186H/186R	186R/186R	Tm:50
w/FSPI	113	113	113		
Control:			90	90	
Uncut = 30			23	23	
Cut = 15					

Table A.2G

APOBEC3F A108S	Uncut	108A/108A	108A/108S	108S/108S	Tm: 54
w/BlpI	279		279	279	
Control:		151	151		
Uncut = 40		128	128		
Cut = 20					

Table A.2H

APOBEC3F 231IV	Uncut	231I/231I	231I/231V	231V/231V	Tm: 52
w/BSmaI	156	156	156		
Control:			130	130	
Uncut = 30			26	26	
Cut = 15					

Table A.2I

APOBEC3D R97C	Uncut	97R/97R	97R/97C	97C/97C	Tm: 53
w/HaeII	180		180	180	
Control:		99	99		
Uncut = 40		81	81		
Cut = 20					

Table A.3

Restriction enzymes used for A3G, A3F, A3D, and A3H restriction digest genotyping assay.

A3 coding region	Restriction Enzyme
APOBEC3H R18L	Fnu4HI
APOBEC3H G105R	HhaI
APOBEC3H K121D	HpyAV
APOBEC3H E178D	SmlI
APOBEC3G H186R	FSPI
APOBEC3F A108S	BlpI
APOBEC3F 231IV	BSmaI
APOBEC3D R97C	HaeII

Table A.4

Synthesized oligonucleotides designed for restriction digest genotyping internal controls. A3G, A3F, and A3D enzyme recognition sites high labeled in GREEN and cut sites are labeled in RED.

SNP amplicon	Primer Sequence	Restriction Enzyme
H186R CON F	AAATTTAAATTTGCGCAAAATTTAAATTT	FSPI
H186R CON R	AAATTTAAATTTTGCGCAAATTTAAATTT	FSPI
A108S CON F	AATTTAAATTTAAATTTGGCTGAGCTTTTAAATTTAAATT	BlpI
A108S CON R	AATTTAAATTTAAAAGCTCAGCCAAATTTAAATTTAAATT	BlpI
A231V CON F	TTTTCACCTGTCTCCTAAATTTAATTTAAAAT	BSmaI
A231V CON R	AAAAGTGGACAGAGGATTTAAATTAAATTTTA	BSmaI
R97C CON F	AAATTTAAATTTAAAGGCGCTAAATTTAAATTTAAATTTT	HaeII
R97C CON R	TTTAAATTTAAATTTCCGCGATTTAAATTTAAATTTAAAA	HaeII
R18L CON F	TTTAAAATTTTAAATTTCGCCGCAAAATTTTAAATTTAAA	Fnu4HI
R18LCON R	AAATTTTAAAATTTAAAGCGGCGTTTTAAAATTTAAATTT	Fnu4HI
G105R CON F	TTTAAAATTTTAAATTTGCGCAAAATTTTAAAATTTAAAA	HhaI
G105R CON R	AAATTTTAAAATTTAAACGCGTTTTAAAATTTTAAATTTT	HhaI
K121D CON F	TTTAAAATTCCTTCTCTCTGTAAATTTTAAAATTTTAAAA	HpyAV
K121D CON R	AAATTTTAAGGAAGAGAGACATTTAAAATTTTAAAATTTT	HpyAV
E178D CON F	TTTAATTTAAATTTAAATTCTTGAGTTTTAAAATTTAAAA	SmlI
E178D CON R	AAATTAAATTTAAATTTAAGAACTCAAAATTTTAAATTTT	SmlI

Table A.5

Nested PCR primers and protocols for near-full-length HIV-1 provirus amplification from donor CD4+ T-cells from Bruner *et al.* **2016** (*87*).

PCR	Length	Primer Name	HXB2 position	Primer sequence	Extension time
Outer PCR	9,064	BLOuterF	623 - 649	AAATCTCTAGCAGTGGCGCCCGAACAG	10 m
		BLOuterR	9,662 - 9,686	TGAGGGATCTCTAGTTACCAGAGTC	
Inner PCRs					
gag	1,448	5GagIn	836 - 857	GGGAAAAAATTCGGTTAAGGCC	1 m 30 s
		3GagIn	2,264 - 2,283	CGAGGGGTCGTTGCCAAAGA	
env	2,841	5EnvIn	6,201 - 6,231	GAGAAAGAGCAGAAGACAGTGGCAATG AGAG	3 m 30 s
		3EnvIn	9,007 - 9,042	CTTGTAAGTCATTGGTCTTAAAGGTACC TGAGGTCTG	
A	4,449	275F	646 - 666	ACAGGGACCTGAAAGCGAAAG	5 m
		3INOut	5,072 - 5,094	AATCCTCATCCTGTCTACTTGCC	
B	5,793	263F	651 - 672	GACCTGAAAGCGAAAGGGAAAC	6 m
		3AccOut	6,421 - 6,443	GGCATGTGTGGCCCARAYATTAT	
C	6,385	5INOut	3,248 - 3,270	ACTCCATCCTGATAAATGGACAG	6 m 30 s
		BLInnerR	9,604 - 9,632	GCACTCAAGGCAAGCTTTATTGAGGCTA	
D	4,778	5AccOut	4,899 - 4,922	CGGGTTTATTACAGGGACARCARA	5 m
		280R	9,650 - 9,676	CTAGTTACCAGAGTCACACAACAGACG	

Outer PCR cycling conditions:
94°C for 2 m; then 94°C for 30 s, 64°C for 30 s, 68°C for 10 m for 3 cycles; 94°C for 30 s, 61°C for 30 s, 68°C for 10 m for 3 cycles; 94°C for 30 s, 58°C for 30 s, 68°C for 10 m for 3 cycles; 94°C for 30 s, 55°C for 30 s, 68°C for 10 m for 21 cycles; then 68°C for 10 m. (30 total cycles)

Inner PCR cycling conditions: (**X** = extension time from the table above)
94°C for 2 m; then 94°C for 30 s, 64°C for 30 s, 68°C for **X m** for 3 cycles; 94°C for 30 s, 61°C for 30 s, 68°C for **X m** for 3 cycles; 94°C for 30 s, 58°C for 30 s, 68°C for **X m** for 3 cycles; 94°C for 30 s, 55°C for 30 s, 68°C for **X m** for 36 cycles; then 68°C for **X m**. (45 total cycles)

Table A.6

Primers and amplification regions for HIV-1 RT used for circulating virus isolation and typing.

Primer Sequences	NAME	ORIGIN	protein	direction	HXB2 start	HXB2 end	Length
GAAGGGCACACAGCCAGAAATTGCAGGG	5'CP1	HIV	RT	F	1981	2008	28
GCCCCTAGGAAAAAGGGCTGTTGG	PRTO5	HIV	RT	F	2008	2031	24
CCTAGGAAAAAGGGCTGTTGGAAATGTGG	2.5	HIV	RT	F	2011	2039	29
CAAACTCCCACTCAGGAATCCA	RT3798R	HIV	RT	R	3798	3777	22
CTAACTGGTACCATAATTTCACTAAGGGAG	2.3	HIV	RT	R	3837	3808	30
GCTCCTACTATGGGTTCTTTCTCTAACTGG	RT3.1	HIV	RT	R	3859	3830	30
CATTGCTCTCCAATTACTGTGATATTTCTCATG	OUT3	HIV	RT	R	4295	4263	33

Sequences provided by Dr. Zabrina Brumme (Simon Fraser University)

Table A.7

Nested PCR primers for near-full-length HIV-1 From Grossmann *et al.* **2015 (*128*)**.

Primer_ID	Seq (5′→3′)	HXB2 position	Ref.
Amplification primers			
$F1_{Gag\text{-}Vpu}$			
0682F	TCTCTCGACGCAGGACTCGGCTTGCTG	0682→0708	[4,21]
0776F	CTAGAAGGAGAGAGAGATGGGTGCGAG	0776→0800	[4]
6352R	GGTACCCCATAATAGACTGTRACCCACAA	6352→6324	[17]
6231R	CTCTCATTGCCACTGTCTTCTGCTC	6231→6207	[17]
$F2_{Vif\text{-}3LTR}$			
5550F	AGARGAYAGATGGAACAAGCCCCAG	5550→5574	[17]
5861F	TGGAAGCATCCRGGAAGTCAGCCT	5861→5884	[17]
9555R	TCTACCTAGAGAGACCCAGTACA	9555→9533	Present study
Sequencing primers			
$F1_{Gag\text{-}Vpu}$			
0776F	CTAGAAGGAGAGAGAGATGGGTGCGAG	0776→0800	[4]
1231F	TCACCTAGAACTTTRAATGCATGGG	1231→1255	Present study
1817F	TAGAAGAAATGATGACAG	1817→1834	[22]
2586F	AAGCCAGGAATGGATGGCCCA	2586→2606	[23]
2713R	GGATTTTCAGGCCCAATTTTTG	2713→2692	Present study
3885R	CTGCTCCATCTACATAGAA	3885→3867	[24]
4350R	CACAGCTAGCTACTATTTCTTTTGC	4350→4326	[24]
4900F	GGGTTTATTACAGGGACAGCAGAG	4900→4923	[25]
5066R	ATCATCACCTGCCATCTGTTTTCCAT	5066→5041	[26]
6231R	CTCTCATTGCCACTGTCTTCTGCTC	6231→6207	[17]
$F2_{Vif\text{-}3LTR}$			
5861F	TGGAAGCATCCRGGAAGTCAGCCT	5861→5884	[17]
6559F	GGGATCAAAGCCTAAAGCCATGTGTAA	6559→6585	[22]
7002F	TTRTTAAATGGTAGTATAGC	7002→7021	Present study
7373R	GAAAAATTCTCCTCYACAATTAAA	7373→7350	Present study
7761F	GTGGGAATAGGAGCTGTGTTCCTTGGG	7761→7787	[22]
8445R	CTCTCTCTCCACCTTCTTCTTC	8445→8424	[22]
9555R	TCTACCTAGAGAGACCCAGTACA	9555→9533	Present study

.

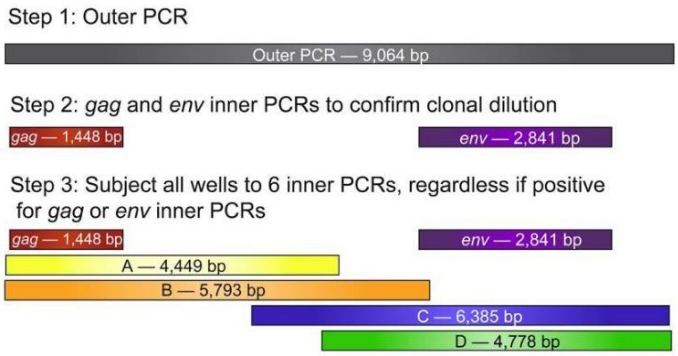


Figure A.1 Nested PCR products for near-full-length HIV-1 provirus amplification from donor CD4+ T-cells. Adapted from Bruner *et al.* **2016 (*87*)**.