Virus-driven evolution of an antiviral gene in deep and shallow time

Alex A. Compton

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2012

Reading Committee:
Michael Emerman, Chair
Jaisri Lingappa
Daniel B. Stetson

Program Authorized to Offer Degree:
Molecular and Cellular Biology

University of Washington

Abstract

Virus-driven evolution of an antiviral gene in deep and shallow time

Alex A. Compton

Chair of the Supervisory Committee:
Affiliate Professor Michael Emerman
Microbiology

The emergence of AIDS in the early 20th century has provoked studies to better understand the evolutionary history of viruses and the factors that govern their spread. Pandemic Human Immunodeficiency Virus-type 1 (HIV-1), which currently infects 34 million people worldwide, emerged following the transmission of a lentivirus between chimpanzees and humans. A growing list of apparently nonpathogenic, species-specific simian strains (known as Simian Immunodeficiency Virus) has now been characterized in dozens of African primates, suggesting that primate lentiviruses are older and more widespread than originally thought. To estimate the extent to which primates and lentiviruses have coexisted, and to determine whether past and present lentivirus infections exhibit pathogenesis, we tracked the interaction between host and virus on a molecular level over evolutionary time. Specifically, we characterized the lentivirus-driven evolution of host restriction factor *APOBEC3G* (*A3G*) in Old World Monkey (OWM) species. We found that residues 128 and 130 of A3G, which determine susceptibility to antagonism by the lentiviral accessory protein Vif, are undergoing recurrent adaptive evolution in both ancestral and contemporary primate populations. We used a broad panel of SIV Vif isolates to demonstrate that natural variation in OWM A3G confers resistance to Vif-mediated degradation, suggesting that adaptive variants of the host factor were selected upon exposure to pathogenic lentiviruses at least 5-6 million years ago (MYA). Furthermore, in members of the divergent *Colobinae* subfamily of OWM, a multi-residue insertion event in A3G that arose approximately 12 MYA blocks the activity of Vif, suggesting an even more ancient origin of SIV. In response to these two adaptive strategies employed by OWM hosts, Vif proteins have counter-evolved to target distinct surfaces of the A3G substrate. Furthermore, some Vif proteins, like that of SIVsm infecting sooty mangabeys, have evolved broad specificity that may facilitate cross-species transmission events. Our findings support that a genetic conflict between primates and lentiviruses has been underway for millions of years and continues to this day. The ancient and ongoing conflict we described may have important implications for our understanding of HIV pathogenesis and spread in human populations. Our studies reveal that, while primate lentiviruses may have modern consequences for human health, they have ancient origins in our non-human primate relatives.

Table of Contents

List of Figures

List of Tables

Acknowledgement

I would like to thank Michael for being a kind, supportive, and inspiring mentor. At the time of writing, he has been selected to receive the FHCRC McDougall mentorship award, an honor that he very much deserves. Over the years, he has crafted a special laboratory atmosphere that has indelibly touched all who have passed through it.

I want to acknowledge Lily for her general expertise, as well as the cadre of former graduate students who taught me most (if not all) that I know: Semih Tareen, Melody Li, Efrem Lim, and Nisha Duggal.

I also have to thank the previous mentors that provided a strong foundation to fuel my scientific ambitions. I am much indebted to Barbara Stebbins-Boaz and Sharon Rose in the Biology department at Willamette University, and Mark Connors and Stephen Migueles of NIAID at the National Institutes of Health.

I am grateful for the scholarships I have received over the last three years. The National Science Foundation Graduate Research Fellowship and the National Institutes of Health Viral Pathogenesis Training Grant provided the financial support for my tuition and research costs, while grants from the UW Center for AIDS and STD Research and the Student-Postdoc Advisory Committee provided funds for travel and conference attendance.

Chapter I: Introduction

Viruses, obligate intracellular parasites, are undoubtedly the most numerous biological entities on Earth. In fact, the ‘virosphere’ likely extends to every cellular species in existence, in every ecological niche. Viruses may have emerged in the RNA world, predating the last universal cellular ancestor (LUCA) from which all life is descended [1]. Alternatively, the ancestry of viruses may lie with host genes that have “escaped” the genomes of cellular organisms and evolved autonomy [2]. The two competing theories may not be mutually exclusive, since some viruses clearly have very recent beginnings. For example, hepatitis delta virus (HDV) encodes a ribozyme sequence found in a human intron and requires human hepatitis B virus to replicate, indicating that HDV was borne out of our genome [3]. While their ultimate origins remain a mystery, it is undeniable that viruses have played a significant role in the evolution of the hosts that they infect.

Retroviruses and mammals

An illustration of the intimate association between eukaryotes and viruses can be gleaned from viral contributions to host genomic content. Endogenous Viral Elements (EVEs), full-length or fragmented viral genomes that have become incorporated into host germ line cells, leave behind a permanent record of past infectious encounters. Initially, these viral artifacts showed relation to exogenous viruses that depend on genome integration for replication, but now EVEs resembling a number of modern virus families have been characterized [4,5].

One virus family that is particularly well represented in animal genomes is *Retroviridae*. Retroviruses are reverse transcribing RNA viruses that integrate into the genome of the host cell as an obligate step of the viral life cycle (Figure 1). The sheer abundance of retrovirus-derived DNA (about 5-8% of the human genome) attests to their success not only as genetic parasites, but also as sources of genetic novelty that have punctuated vertebrate evolution [5]. Exapted retroviral genes (integrated virus-derived elements that have been repurposed for a role in the host cell) are responsible for major innovations of mammalian reproduction, such as formation of the placenta and the transition to life births [6,7].

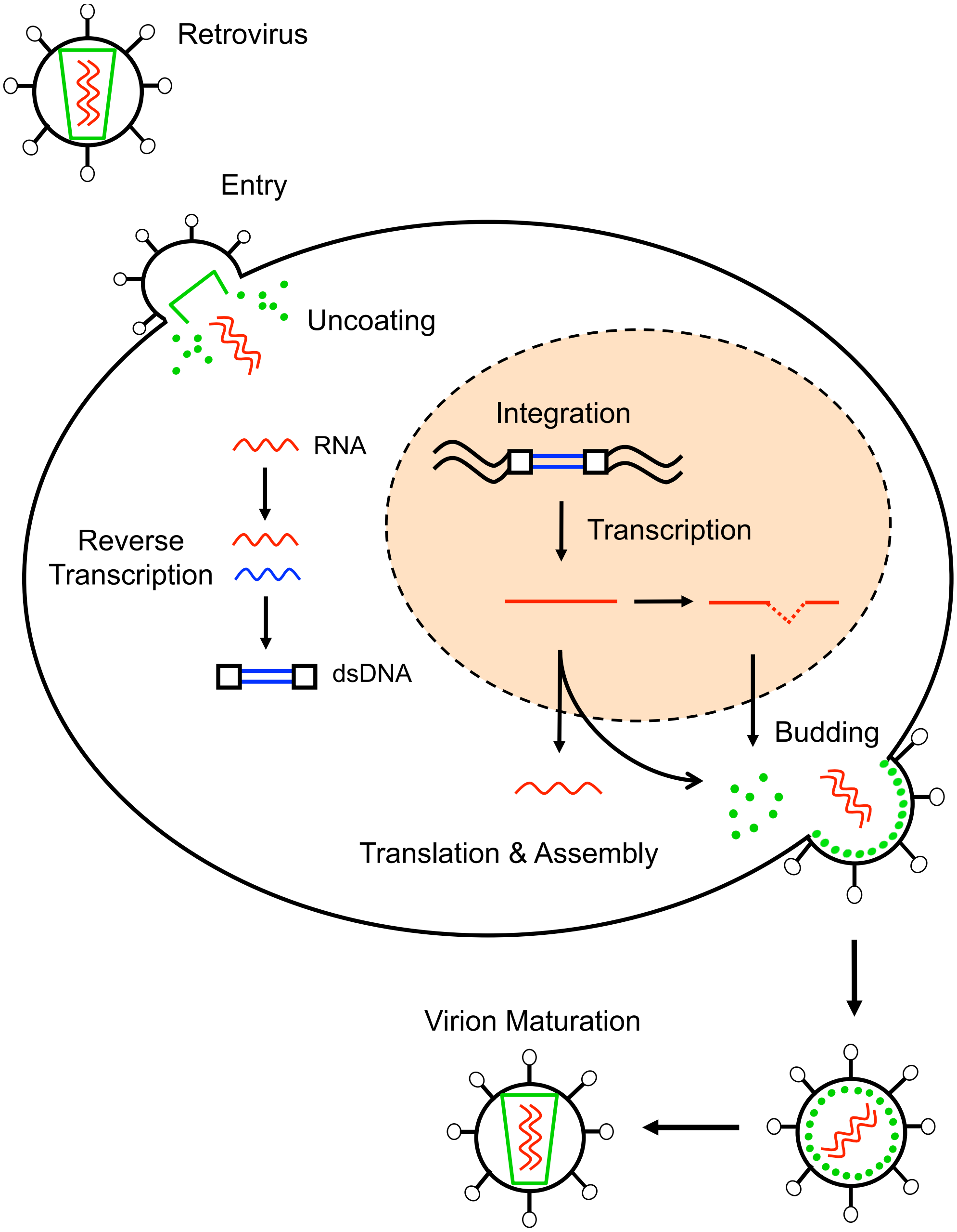


Figure 1. The retrovirus life cycle
Shown is a simplified diagram of retrovirus replication in a host cell. An enveloped virion containing capsid protein, two copies of the ssRNA genome, and pre-synthesized viral and host factors attaches to the host cell via specific interaction with a membrane-associated receptor. Fusion of the viral-associated and target cell membrane ensues, followed by breakdown of the capsid coat. Reverse transcriptase produces a DNA copy of the viral genome that is imported into the nucleus for integration into the host chromosome (provirus). Proviral DNA is transcribed by cellular RNA polymerase, producing full-length viral genomes and viral mRNAs needed for virion manufacture. Nascent virions egress and undergo maturation before becoming infectious.

Non-human primates: retrovirus reservoirs with zoonotic potential

Decades of research devoted to the role of oncogenic retroviruses in animals preceded the discovery of human retroviruses as causative agents of disease. Tracing the sources of *Retroviridae* infections in humans has uncovered an immense reservoir of related viruses in African non-human primates. Deltaretroviruses, spumaviruses, and lentiviruses epitomize the profound potential for cross-species transmission that retroviruses exhibit, a property that has led to the rise of human-specific viruses. While the molecular determinants for zoonotic transmission are complex and incompletely understood, the common thread underlying retrovirus emergence in humans is bushmeat hunting in west central Africa. The killing and consumption of primates, including both monkeys and ape species, has increased over the last few decades as a result of urban development and commercial logging practices. The construction of roads into remote forests have enabled hunters to penetrate previously inaccessible habitats and transport bushmeat back to city markets, increasing the contact rate between humans and non-human primates [8].

Primate T-lymphotropic viruses (PTLVs) of the Deltaretrovirus genus cause chronic infections in humans (termed HTLV) and in more than 30 simian species in sub-Saharan Africa (STLV) [9]. Four strains of HTLV (types I, II, III, and IV) have been characterized to date, with three showing clear phylogeographic relatedness to STLV strains endemic to Old World Monkeys (OWM) (see Figure 2 for a primate phylogeny). That is, viral nucleic acid sequences from human infections cluster in a phylogenetic tree with viruses isolated from OWM in proximal habitats, indicative of simian-to-human cross-species transmission. Moreover, the spatial distribution of genetically similar STLVs indicates that this virus family is spread predominantly via interspecies virus transmission [10]. The majority of HTLV infections appear to be asymptomatic, but approximately 5% of cases are associated with T-cell leukemia or lymphoma, or an inflammatory syndrome of the central nervous system [10]. Similarly, no pathology has been observed among infected non-human primates, except for some instances of lymphoma in captive gorillas [11].

In contrast to PTLVs, Simian Foamy Viruses (SFVs, of the Spumavirus genus) have been identified at high prevalence in a wide variety of primates including OWM, New World Monkeys (NWM), and even prosimian species of Madagascar. With the exception of humans, nearly every primate species may harbor a species-specific strain of SFV. There is no disease associated with SFV infection, and phylogenetic clustering of virus strains mirrors that of the primate host phylogeny, suggestive of host-virus coevolution for more than 60 million years [12]. Nonetheless, new SFV infections may also result from cross-species transfer. SFV strains from a number of monkey and ape species have been reported in humans exposed by occupation (pet owners, temple workers, and hunters), giving rise to chronic viral infection that persists for decades with no apparent accompanying disease [13,14]. However, there is no human foamy virus designation because so far there is no evidence of secondary human-to-human transmission [15,16].

While the frequent transfer of STLV and SFV into humans underlines the effects of constant exposure to non-human primates, neither has achieved the pandemic status attained by lentiviruses. An appreciation for the nature of retrovirus infections in non-human primates will help us understand what drives the pathogenic outcome of HIV-1 infection in humans.

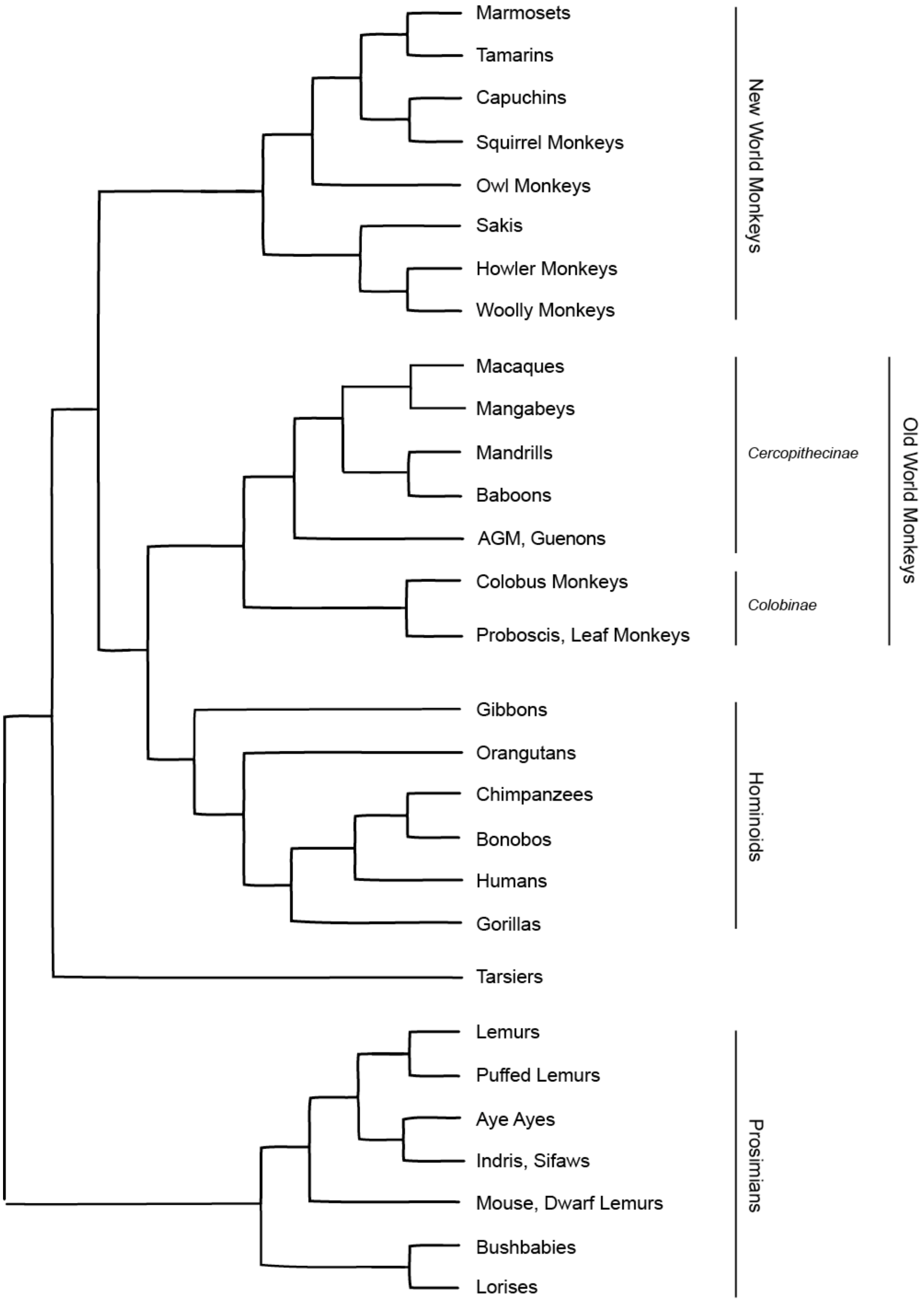


Figure 2. A species phylogeny of selected primates
The evolutionary relationship of living primates is shown, with branching orders reflecting the currently accepted species phylogeny [17]. The Old World Monkeys (OWM) are subdivided into the *Cercopithecinae* and *Colobinae* subfamilies.

Origins of Human Immunodeficiency Virus

The first known lentivirus infections of humans were recognized in 1983, two years after presentation of rare opportunistic infections and malignancies in American homosexual men. Although initially misidentified as a deltaretrovirus (HTLV-III) [18,19], the retrovirus responsible shared similarities to animal lentiviruses Equine Infectious Anemia Virus (EIAV) and Maedi-Visna Virus [20], and it was eventually dubbed Human Immunodeficiency Virus (HIV). The subsequent discovery of a second AIDS-causing virus in western Africa [21] hinted that HIV was derived from a non-human primate source. The second virus, named HIV type 2 (HIV-2), was only distantly related to HIV-1 but bore semblance to a simian virus that caused immunodeficiency in captive macaques (Simian Immunodeficiency Virus, SIV) [22,23]. Characterization of the SIV reservoir identified that simian relatives of HIV-1 and HIV-2 were found at high prevalence in chimpanzees and sooty mangabeys, respectively, providing the first evidence that the emergence of AIDS in humans resulted from cross-species transmission of lentiviruses from different non-human primate species [24,25]. A phylogeographic approach, combining sequence analysis with the spatial distribution of naturally circulating primate lentivirus, demonstrated unequivocally that HIV-1 and HIV-2 arose from multiple zoonotic transfers from African primates [26-28] (Figure 3).

Global HIV-1 diversity falls into four distinct lineages or groups (M, N, O, and P), each the product of an independent cross-species transmission event. The first to be discovered, group M (main) is the pandemic strain that accounts for most infections worldwide. Group O and group N account for less than 1% of infections, while group P has been identified in just two unrelated individuals [29,30]. Phylogenetic analysis of HIV-1 sequences, time-calibrated by the viral mutation rate (molecular clock), revealed that the chimp-to-human transmission resulting in group M occurred at the turn of the 20th century [31], while the events leading to other groups occurred more recently. Groups M and N of HIV-1 are very closely related to SIV infecting a chimpanzee subspecies (*Pan troglodytes troglodytes*) in southern Cameroon. The source of group O is less certain because no known ape viruses are particularly reminiscent, while the single isolate of group P draws ancestry from SIVgor of gorilla origin (itself the descendant of SIVcpz) (Figure 3).

As opposed to pandemic HIV-1, which has infected an estimated 60 million people to date, HIV-2 remains largely confined to West Africa [32]. Viral loads in HIV-2 infected individuals tend to be lower than those observed in HIV-1 infected persons, which may explain its lower transmissibility. In fact, many HIV-2 infected individuals appear to spontaneously control viral replication, with only a minor subset developing AIDS-like symptoms [33]. Sequence analysis of HIV-2 identified the virus reservoir as the sooty mangabey (*Cercocebus atys*), an OWM that harbors a lentivirus (SIVsm) circulating in the vicinity of infected individuals [34]. To date, eight distinct lineages of HIV-2 have been named, each deriving from an independent host transfer. In accordance with the HIV-1 nomenclature, the groups were named A-H, although only groups A and B may be capable of extensive secondary spread [29] (Figure 3). Therefore, encounters between apes and monkeys are responsible for the surfacing of HIV, and because primate lentiviruses are naturally transmitted via infected blood and body fluids, humans must have been exposed in a similar fashion. As is likely the case for HTLV and SFV infections in humans, bushmeat hunting in west central Africa is the most likely scenario whereby lentiviruses entered human populations during the last century.

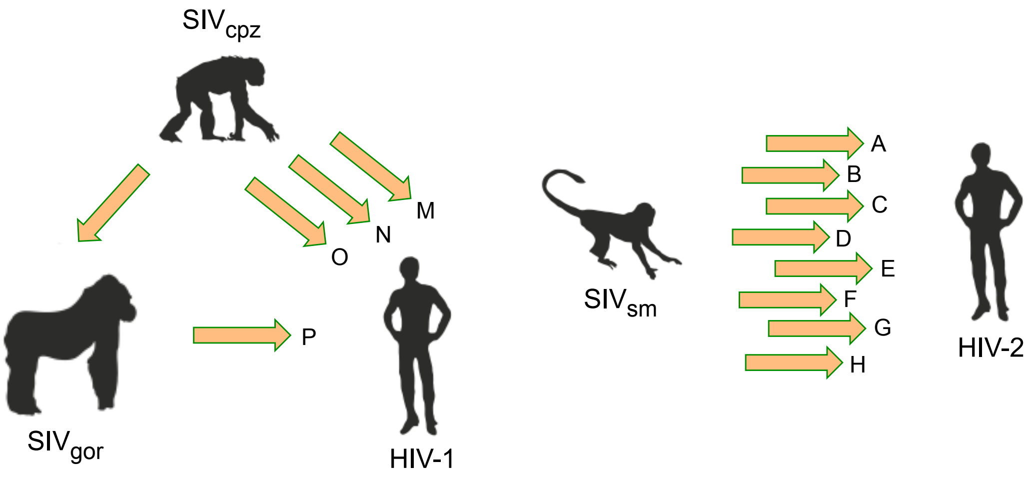


Figure 3. Human lentivirus infections originated from multiple cross-species transmissions of SIV
Left: at least three transmissions of SIVcpz from chimpanzees has given rise to HIV-1 groups M, N, and O, while a single case (group P) is derived from SIVgor in humans. SIVgor itself is the product of cross-species transmission of SIVcpz. Right: at least eight transmissions of SIVsm from sooty mangabeys has given rise to HIV-2 groups A-H, although only groups A and B appear capable of substantial secondary spread.

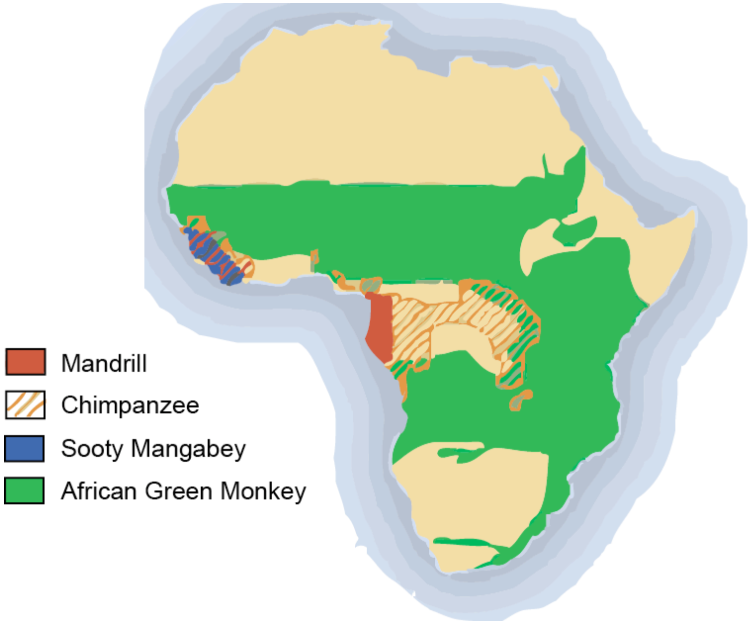


Figure 4. The geographic distribution of select non-human primates harboring SIV
Shown are the ranges of chimpanzees (orange stripes) and sooty mangabeys (blue), the zoonotic reservoirs of HIV-1 and HIV-2, respectively. Also shown is the wide distribution covered by the most abundant non-human primate species, AGM (green). The range of mandrills is shown for comparison (red). Adapted and reprinted with permission from [35] under license number 3032750924903.

The SIV reservoir in non-human primates

Serological evidence of lentivirus infection suggests that over 40 primate species are associated with SIV in the wild, and viral isolates have been obtained from the majority of these [36]. Molecular characterization demonstrates that most primates harbor a single strain of SIV that is species-specific—viral sequences obtained from the same species form a monophyletic clade (share a most recent common ancestor).

Ongoing work in the field of primate genetics has produced a robust phylogeny of living primate species, reconstructing the evolutionary relationship between them and providing dates of divergence from ancestral populations [17]. A comparison of the primate species phylogeny with that of the lentivirus pool reveals that some virus lineages may have been associated their respective hosts for a very long time. For example, SIVs infecting four closely related species of African green monkeys (AGM, *Chlorocebus* species) exhibit common ancestry, suggesting that, like SFV infections, SIV may have coevolved with their respective hosts [37,38], although this has been disputed [39]. However, there is substantial evidence that some SIV strains have jumped species, leading to discordant topologies of virus and host phylogenetic trees (Figure 5). The outcome of cross-infections can vary, ranging from "dead end" incidental infections [40] to the generation of novel lentivirus lineages capable of secondary spread and persistence in the new host population [41]. Furthermore, the identification of circulating recombinant forms of SIV demonstrates that co-infection by two divergent lentivirus strains can occur. In fact, some primate species, like mandrills (*Mandrillus sphinx*) and mustached guenons (*Cercopithecus cephus*), harbor two lentivirus strains as a result of recombination between primary and secondary SIV infections [42,43]. Thus, in addition to long-standing host-virus relationships, some natural SIV infections have emerged more recently through cross-species transmission and recombination. In contrast to SFV infections, the prevalence of SIV infections varies substantially, ranging from 1% in some species to over 50% in others [44]. Variable seroprevalence may be a reflection of differential host-virus associations. That is, SIV infections achieving a high seroprevalence are suggestive of cospeciation while those that are less prevalent likely emerged relatively recently via cross-species transmission.

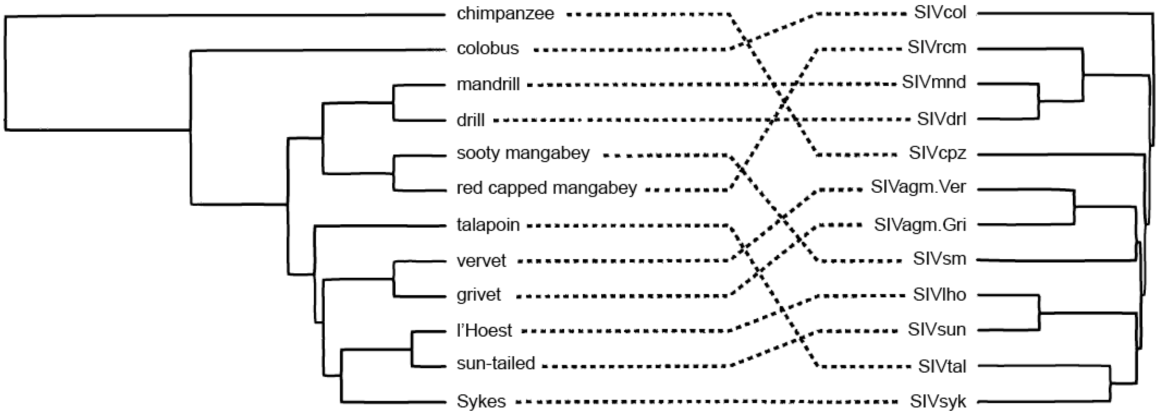


Figure 5. Discordance between host and virus phylogenies reveal cross-species transmission events
A species phylogeny of select primates is juxtaposed with phylogenetic analysis of SIV *pol* genes. Pairings between host and virus produces a tanglegram, whereby non-horizontal dotted lines indicate phylogenetic discordance and, hence, cross-species transmission. This analysis reveals that, while some lentiviral lineages may have co-evolved with their respective hosts, some must be younger by comparison. Reprinted with permission from [45] under license number 3032751164056.

In either case, the association between natural hosts and SIV is thought to be considerably older than that of humans and HIV [46,47]. Pandemic HIV-1 infections have claimed an estimated 25 million lives, and around 33 million people are currently infected [48]. However, natural SIV infections do not generally cause an AIDS-like immunodeficiency in their autologous host species, leading to the hypothesis that the virus-host relationship has evolved towards an apathogenic state [49-52]. In stark contrast to the lentivirus infections plaguing humans and macaques, SIV infection of some African primates appears to induce an asymptomatic, nonprogressive chronic infection [35]. Two 'natural hosts' of SIV, sooty mangabeys and AGM, have been the subject of intense study due to their abundance and availability in primate centers in the United States and Europe.

In efforts aiming to characterize the mechanisms underlying the lack of disease associations, it was found that SIV infections in these monkeys achieve viral loads similar to those observed in HIV-1-infected humans [50,53,54]. Furthermore, they mount adaptive immune responses in response to SIV infection that are unable to fully suppress viral replication, demonstrating that lack of disease is not the result of immune-mediated control [55]. However, a growing list of immunomodulatory phenotypes have been identified in sooty mangabeys and AGM, which collectively may serve to protect them from lentivirus-induced AIDS. These traits are noticeably absent from 'nonnatural hosts,' mainly humans and macaques. Some key features specific to the infection of natural hosts are: 1) preservation of healthy CD4+ T cell counts in peripheral circulation; 2) preservation of the central memory subset of CD4+ T cells; 3) preservation of mucosal immune architecture and absence of microbial translocation; and 4) lack of chronic immune activation [35] (Figure 6). Most of these adaptive differences are associated with resolution of the innate immune response initiated during primary infection, a protective phenotype that will be difficult to induce in humans with traditional vaccine-based approaches. Nonetheless, the genetic basis for immunomodulatory adaptations will be invaluable to our basic understanding of HIV-1 pathogenesis, and will aid in the development of novel anti-HIV therapies.

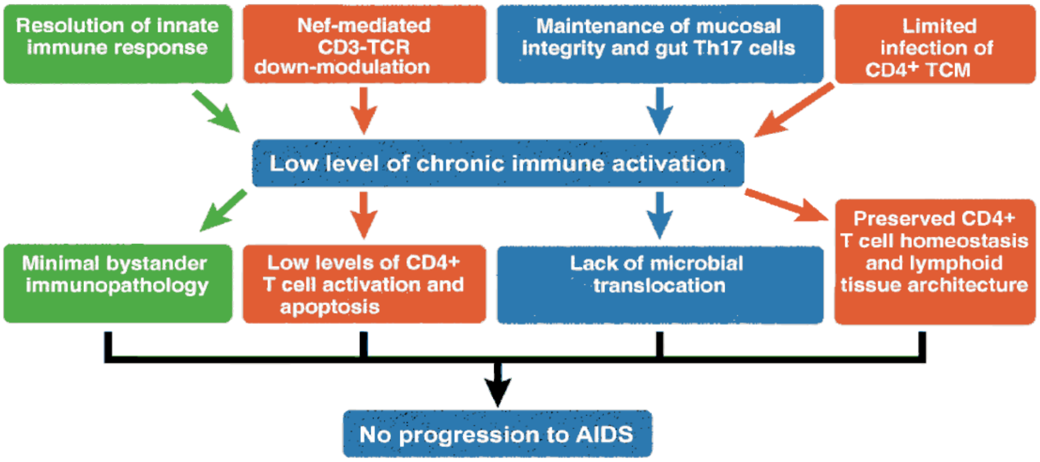


Figure 6. Immunomodulatory phenotypes exhibited by natural hosts of SIV infection
Viral and host factors associated with the lack of AIDS progression in AGM and sooty mangabeys are shown. A protective phenotype is thought to result from multiple genetic adaptations, each acting as a cause or consequence of decreased levels of chronic immune activation. Adapted and reprinted with permission from [35] with license number 3032750924903.

Despite the lack of overt AIDS in sooty mangabeys and AGM, the age and pathogenic potential of wild SIV infections in diverse primate taxa remain largely uncharacterized. Initial attempts at calculating the age of SIV using phylogenetics produced widely disparate, but all relatively recent, estimates [56,57]. However, two significant findings have pushed back considerably the age estimates of primate lentiviruses. First, the discovery of a full-length endogenous SIV in the genomes of Malagasy lemurs indicates that lentiviruses were present in prosimians at least 4 million years ago (MYA) [58,59]. Second, a date to calibrate SIV phylogenetics was made possible with the identification of SIV strains endemic to the African island of Bioko.

Here, each virus found on the island shares ancestry with a mainland virus, and their respective hosts belong to the same genus, demonstrating that lentiviruses have been infecting OWM for at least 12,000 years, and probably much longer [60]. The use of viral sequences to establish the age of virus families is problematic because rapid evolution obscures phylogenetic signals, and because many viral lineages have gone extinct in the past [61,62]. As an alternative strategy to dating the extent to which primates have been associated with lentivirus infections, we assess how lentiviruses have influenced the evolution of host immunity.

Cell-intrinsic innate immunity and virus-driven evolution

Like all known viruses, lentiviruses depend on numerous cellular factors and pathways to complete the viral life cycle and spread to new target cells. Known as proviral factors or virus cofactors, these are host molecules that are commandeered by the virus in order to infect the cell [63]. To prevent virus takeover, mammalian cells also express a battery of antiviral effectors, known as host restriction factors, that function to block virus replication at several stages [64]. In response, viruses have evolved counter-measures that allow evasion from or direct antagonism of these cellular proteins [65]. However, as their name implies, restriction factors are often highly divergent between species, so they may restrict the host tropism of species-specific virus infections *in vitro* and *in vivo*. The establishment of a hostile intracellular environment and the disarming of host defenses by the infecting virus forms the basis of a genetic conflict [66]. In this scenario, host proteins imposing blocks to viral replication are antagonized by viral antagonists, applying a selective pressure to host populations to evolve and evade antagonism. This in turn pressures the virus to counter-evolve, adjusting the nature of the antagonist in order to re-establish interaction with the host factor (Figure 7). Lentivirus genomes evolve one million times faster than those of their hosts [67], and so viral adaptation occurs relatively quickly. Nonetheless, genes encoding host proteins that are critical to the maintenance of the antiviral state will evolve rapidly and recurrently relative to the rest of the genome, displaying signatures of adaptive evolution on surfaces in close contact with virus proteins.

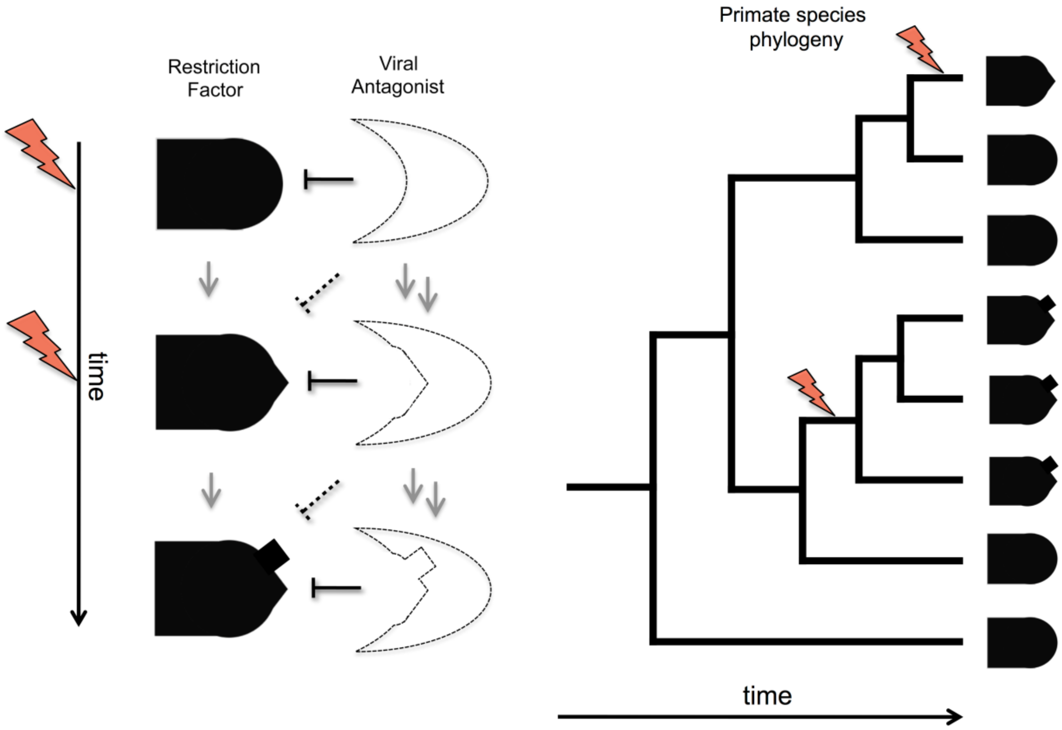


Figure 7. Genetic conflict between a cellular restriction factor and a viral antagonist
Left: A cycle of host-virus coevolution is initiated when a host restriction factor is counteracted by an antagonist encoded by a pathogenic virus. Selection of host factor variants that successfully evade antagonism by the virus is followed by counter-selection of viral variants that re-establish the host-virus interaction, applying selective pressure to the host factor once again. Right: The age of host genetic adaptations, and of the pathogenic viruses responsible, can be estimated when superimposed onto the consensus primate species phylogeny.

Three host restriction factors epitomize the intracellular barriers imposed by innate immunity to lentivirus replication, as well as the processes by which lentiviruses have evolved to overcome them.

TRIM5α

The alpha isoform of the tripartite motif (TRIM)-containing protein TRIM5 is a cytosolic antiviral protein that inhibits infection after retroviral capsids enter target cells. TRIM5α binds directly to the retrovirus capsids to accelerate capsid fragmentation and disrupt the events leading to reverse transcription [68]. Retroviruses can evolve resistance to TRIM-mediated restriction through mutation of the retroviral capsid (CA) protein, encoded by the *gag*, which renders it unrecognizable by the host factor (Figure 8). Consistent with its role as a cellular sentry for invading retrovirus particles, TRIM5α is undergoing positive selection (adaptive evolution) in primates in order to re-establish capsid interactions over evolutionary time [69]. This host-virus dynamic, which generally leaves TRIM5α out of step with its retroviral capsid partner, means that TRIM5α is ineffective against retroviruses infecting the same species yet potently blocking viruses from other species. The role of TRIM5α as inhibitor of cross-species lentivirus transmission was demonstrated *in vivo* in SIV infected macaques. The TRIM5 locus is highly polymorphic in rhesus macaques, generating variants with differential activity (specificity) towards retroviruses (including lentiviruses) [70]. A retrospective analysis of experimental cross-species transmissions of SIV into rhesus macaques found that *TRIM5* genotype of the recipient host largely dictates levels of virus replication observed following transfer. Furthermore, viral adaptation of the capsid gene occurred over time such that TRIM5-resistant virus emerged, uncovering the molecular events that gave rise to the AIDS epidemic in macaque species [71].

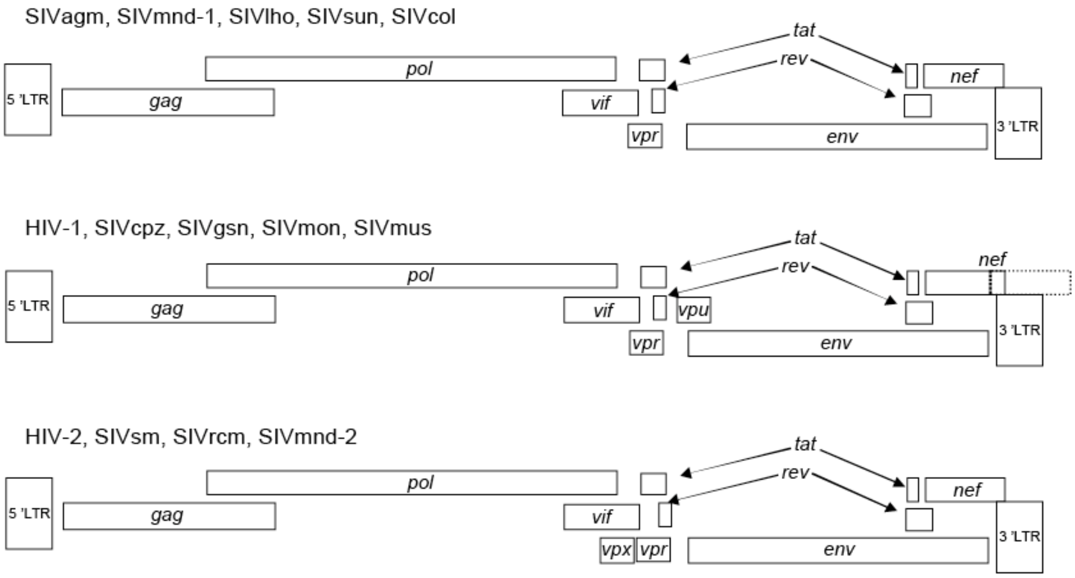


Figure 8. Genomic architecture of primate lentiviruses
Eight lentivirus families have been described to date, and all fall into three schemes of genomic organization. The lineage containing HIV-1 and its predecessor SIVcpz is distinguished by the *vpu* accessory gene, while the HIV-2/SIVsm lineage contains the *vpx* accessory gene, a paralog of the *vpr* gene found in all primate lentiviruses. Other significant differences include configuration of the *nef* gene, which typically overlaps with *env* but not in the case of HIV-1 and SIVcpz. Adapted and reprinted with permission from [36].

Tetherin

The identification of some host restriction factors has been based on findings that lentiviral accessory proteins are required for efficient completion of the viral life cycle. In the absence of the lentiviral accessory protein Vpu (Figure 8), nascent HIV-1 virions were found to adhere to the cell surface of infected cells, suggesting a specific block to virion release [72]. The culprit, a membrane anchored protein called tetherin [73,74], associates with host lipid membranes via a single-pass transmembrane domain as well as a glyco-phosphatidylinositol (GPI) anchor, allowing it to act against a number of disparate virus families that share no sequence or structural homology [75-77]. While originally identified as the target of HIV-1 Vpu, tetherin is deeply mired in genetic conflict with several lentiviral genes that encode antagonists. Most SIV strains do not encode Vpu. Instead, accessory protein Nef fulfills the role as tetherin antagonist in non-human species [78-81]. Remarkably, following the transmission of SIV from chimpanzees to humans, the tetherin antagonist switched from Nef to Vpu in the emerging HIV-1 [80,81]. Notably, Vpu from group M HIV-1 is unique in its ability to downregulate tetherin, suggesting that antagonism of tetherin may be partially responsible for its pandemic potential [80]. Highlighting the extreme adaptability of lentiviruses invading new hosts, HIV-2 has coopted the envelope glycoprotein, Env, to perform anti-tetherin functions [80,81]. The direct antagonism of tetherin by virus accessory proteins fundamentally changes the nature of its adaptive evolution, and demonstrates that the host and virus sit at opposite sides of the table in this genetic conflict. In contrast to TRIM5α, which is evolving rapidly to re-acquire a host-virus interaction, tetherin evolves in order to evade them (Figure 7).

APOBEC3G

The host restriction factor APOBEC3G (A3G) was first identified as being sufficient to prevent productive HIV-1 infection in the absence of the HIV-1 accessory protein Virus infectivity factor (Vif) [82]. As the name of its antagonist implies, A3G saps infectivity from virions produced in its presence. A3G is a cytidine deaminase that becomes packaged into newly assembled lentivirus particles and executes lethal mutagenesis of viral transcripts during the subsequent round of infection. A3G transits insides virions until infection of a new target cell, where it deaminates cytidine nucleotides in the minus strand viral DNA produced during reverse transcription, resulting in G-to-A hypermutation of viral cDNAs [83]. A3G also inhibits virus replication in a deaminase-independent manner by blocking the act of reverse transcription itself [82-84].

A3G is one member of an *APOBEC3* gene cluster located on human chromosome 22, which includes seven paralogs derived from tandem gene duplication events [85]. Vif reverses the restriction imposed by A3G, as well as other family members A3C, A3DE and A3F, by linking them to a cellular E3 ubiquitin ligase complex and accelerating their turnover at the proteasome [86]. However, A3G exhibits broad activity *in vitro* against a panel of retroviruses and other viruses that utilize a single stranded DNA intermediate, an observation that is likely relevant to *in vivo* infections because many viruses have devised strategies to avoid its potent restriction [87-89]. Therefore, the evolutionary pressures shaping *A3G* evolution may originate from multiple sources.

A3G is one of the most rapidly evolving genes in the primate genome, and has been undergoing positive selection (adaptive evolution) for at least 33 million years over the course of evolution. Initial analyses of the gene's evolutionary history exposed signals of selection throughout its length, indicating that many surfaces of the enzyme may be engaged in genetic conflict with viral antagonists [90,91]. However, contributions to host evolution by specific virus families can be dissected with experimental insight.

Antagonism of primate A3G by lentiviral Vif can occur in a species-specific manner, suggesting that A3G orthologs have diverged over the course of primate evolution in a manner that affects antagonism by Vif. Indeed, A3G orthologs from different species exhibit differential susceptibility to the same Vif protein. For example, human A3G is sensitive to antagonism by HIV-1 Vif, but not Vif from SIV infecting AGM, while AGM A3G is antagonized by SIV Vif but not HIV-1 Vif. Divergence of a single amino acid at residue 128 of A3G was shown to underlie this difference in Vif sensitivity, identifying ^{126}FWKPDYQ^{132} as major determinants of the interaction with Vif [92-96]. However, given the high levels of genetic divergence in general between hominoids and OWM, it in unclear whether substitutions in this motif were driven by Vif-mediated selective pressure. A comprehensive examination of *A3G* evolution in primates, taking into account both interspecies divergence and intraspecies diversity, will provide a better view of the impact that lentiviruses have had on primate evolution.

Inferring past infections from virus-driven evolution of host immunity

It is possible to estimate the true evolutionary age of viruses by tracking and dating the evolution of antiviral genes in hosts that they infect [97]. Moreover, if the virus encodes an antagonist of the restriction factor, and the interactive interface between host and viral factors is known, then evolution at the site(s) of interaction can be used to infer past instances of infection [98].

Vif is most likely the oldest lentiviral accessory gene, as it is encoded by all but one member of the *Lentiviridae* genus [47,99], including two endogenous lentiviruses found in animal genomes (Figure 9). The early birth and ongoing retention of *vif* suggest that antagonism of A3G is crucial to lentivirus spread and survival. **Therefore, *A3G* is a likely substrate for signatures of lentivirus-driven selection, from which a detailed account of past viral challenges can be reconstructed.** Using a combination of evolutionary genetics and experimental virology, the objective of my doctoral research was to track the dynamics of the A3G-Vif interaction over the course of primate evolution. The rationale for this approach was three-fold: 1) An understanding of the surfaces of A3G that are rapidly evolving in response to Vif-mediated antagonism will reveal the molecular interface formed with the viral antagonist, knowledge which may be exploited for novel anti-HIV therapeutic interventions; 2) By anchoring instances of lentivirus-driven host evolution into the phylogenetic framework of modern primate species, instances of past lentiviral infections can be reconstructed and an age estimate can be made for the primate lentivirus family; 3) As it is generally believed that most SIV infections occurring in the wild result in an asymptomatic chronic infection, instances of lentivirus-driven evolution in the past will reveal that the immediate ancestors of modern lentiviruses were pathogenic in nature. Similarly, cases of recent selection in host genes may suggest that modern lentiviruses can impact host fitness as well.

My study of the evolutionary dynamics of the A3G-Vif interaction in primates has revealed an ongoing genetic conflict. A3G is undergoing selection in multiple host lineages in a manner that is consistent with Vif-driven evolution, allowing for the inference of pathogenic lentiviruses over time.

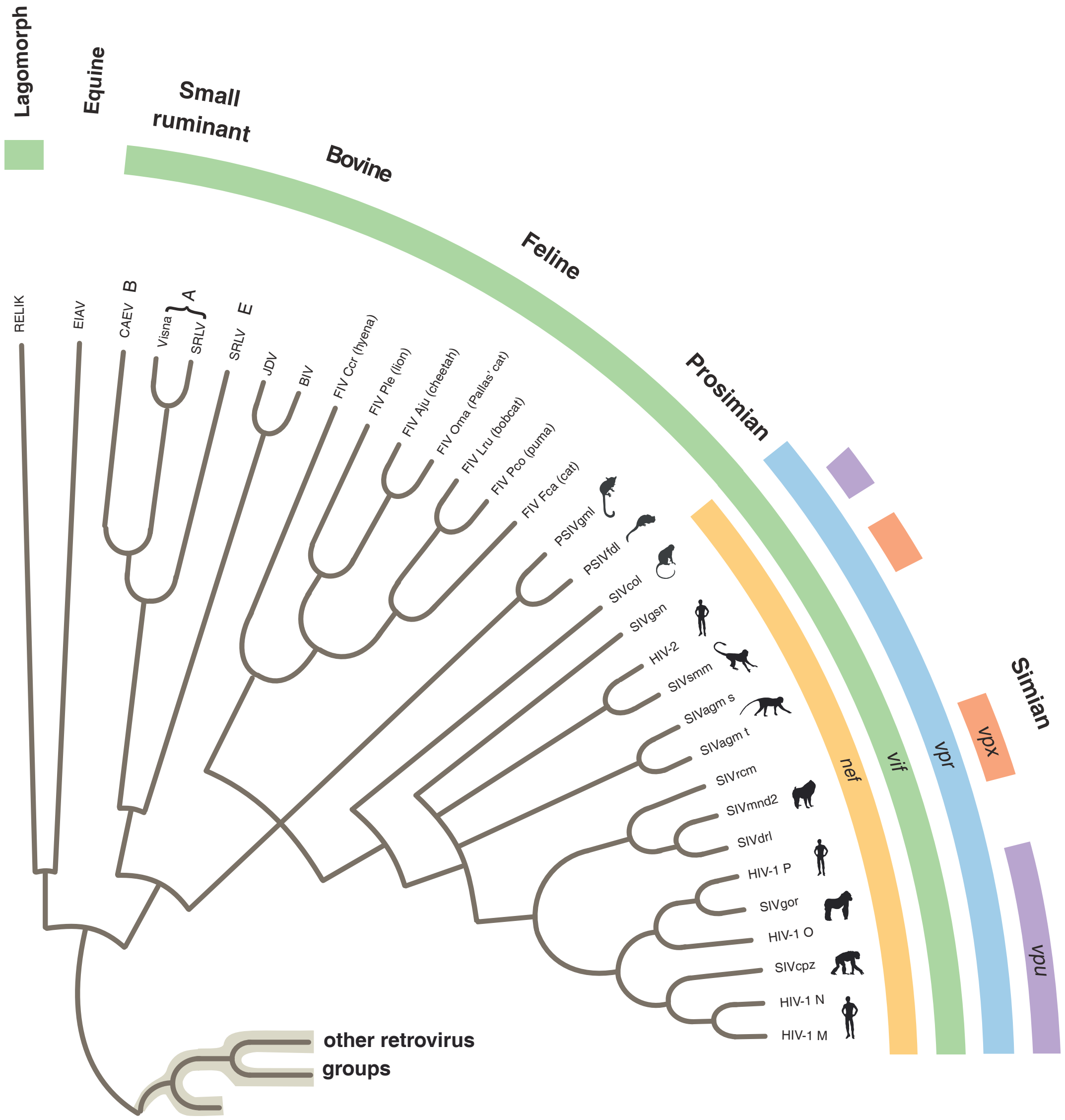


Figure 9. *Vif* is found in nearly all mammalian lentiviruses
Lentiviruses typically encode a number of accessory genes in addition to the fundamental retroviral genes *gag*, *pol*, and *env*. The presence of accessory genes in various lentiviruses is indicated by corresponding bars along the tips of the tree. Not shown are the lentiviral regulatory genes, *rev* and *tat*. Virus names: BIV, bovine immunodeficiency virus; CAEV, caprine arthritis encephalitis virus; EIAV, equine infectious anemia virus; FIV, feline immunodeficiency virus; JDV, jembrana disease virus; PSIV, prosimian immunodeficiency virus; RELIK, rabbit endogenous lentivirus K; SRLV, small ruminant lentivirus. Adapted and reprinted from [47] under license 3032750924903.

Chapter II: Methods

Amplification of Primate Genes and Generation of Expression Plasmids

Sources of AGM samples are described in Figure. Whole RNA and genomic DNA (gDNA) were extracted from AGM peripheral blood mononuclear cells or AGM-derived continuous cell lines using the RNeasy Mini Kit (Qiagen) and the Whole Blood DNA Isolation Kit (Qiagen), respectively. The following fibroblast or lymphoid cell lines derived from additional primate species were obtained from Coriell Cell Repositories (Camden, NJ): patas monkey (*Erythrocebus patas*; AG06116A), mustached guenon (*Cercopithecus cephus*; PR00527), lesser white nosed monkey (*Cercopithecus petaurista*; PR00949), Wolf's guenon (*Cercopithecus wolfi*; PR01241), De Brazza's monkey (*Cercopithecus neglectus*; PR01144), Allen's swamp monkey (*Allenopithicus nigroviridis*; PR01231), red capped mangabey (*Cercocebus torquatus*; PR00485), sooty mangabey (*Cercocebus atys*; G077), Francois' leaf monkey (*Trachypithicus francoisi*; PR01099), proboscis monkey (*Nasalis lavartus*; PR00674), and the mantled colobus monkey (*Colobus guereza*; PR00980). The following transformed lymphoid cell lines were obtained from the NIH Nonhuman Primate Reagent Resource: Olive baboon (*Pabio anubis*, GAG-LCL). Whole RNA was extracted using the RNeasy Mini Kit (QIAGEN). Full-length *A3G* was amplified via one-step RT-PCR with the SuperScript III Reverse-Transcriptase Kit (Invitrogen) using primers specific for OWM A3G (FOR 5'-ATG AAG

CCT CAA ATC AGA AAC ATG G-3’, REV 5’-CAG TTT CCC TGA TTC TGG-3’). Bulk PCR product was subcloned, and six to ten clones were sequenced. If two distinct *A3G* sequences were detected, the individual was considered to be heterozygous. When samples were limited to gDNA only, *A3G* coding exons (1-8) were amplified and sequenced individually using primers targeted to non-coding regions of the locus. Full-length reconstruction of haplotypes originating from grivet monkeys (hap V and VII) was performed by mutagenesis of other AGM *A3G* haplotypes using the Quikchange Site-Directed Mutagenesis Kit (Stratagene). Site-directed mutagenesis was used to generate the E128K and H130D *A3G* mutants. OWM *A3G* sequences were appended with a 5’ hemagglutinin (HA) tag by PCR and cloned into the mammalian expression vector pcDNA3.1. Partial mitochondrial genes *Cytochrome B* and *NADH1* were PCR amplified from AGM gDNA and bulk PCR products were sequenced.

Amplification of SIV *vif* and Production of Recombinant HIV-1

Vif genes were PCR-amplified from full-length molecular clones SIVsab-1 and SIVtan.1 and from plasma derived from monkeys infected with SIVgri.667 and SIVver90. In the latter cases, viral RNA was isolated using the QIAamp Viral RNA Mini Kit (Qiagen) and used as template in Two-Step RT-PCR with *vif*-specific primers containing Mlu1 and Xba1 restriction sites. SIVver90 *vif* derived from experimentally infected sabaeus monkeys was amplified using a nested RT-PCR approach. One to two mL of plasma was concentrated by centrifugation for 30 min at 16,000 rpm. Viral RNA was isolated using the QIAamp Viral RNA Mini Kit (Qiagen). Using specific primers that were targeted to SIVagm.Ver90 *integrase* and *vpr*, *vif* was amplified using a one-step

RT-PCR. Products were amplified in a second round of PCR using nested primers containing Mlu1 and Xba1 restriction sites. The following SIV *vif* sequences were synthesized by GenScript (without codon optimization): SIVdeb CM5, SIVmus-1 CM1085, SIVsm E041, SIVcol CGU1, SIVwrc 98CI04, SIVolc 97CI12. HIV-2_{ROD9} *vif* and SIVmac239 *vif* were PCR amplified from the full-length molecular clone [23,100] and other *vif* genes were previously described in [98]. SIV *vif* sequences were appended with a 5' Kozac sequence and 5'Mlu1 and 3' Xba1 restriction sites by PCR and cloned into the HIV-1Δ*vif* molecular clone pLaiΔ*env*Luc2Δ*vif*, generated after Nde1-Stu1 deletion in pLaiΔ*env*Luc2. Epitope-tagged versions of select *vif* isolates (SIVagm.Sab *vif*, SIVcol *vif,* SIVwrc *vif,* and SIVolc *vif,*) were produced by appending a 3' 3X-FLAG, and they too were cloned into the HIV-1Δ*vif* molecular clone pLaiΔ*env*Luc2Δ*vif*. The resulting proviral plasmids lack *env*, have a firefly luciferase gene inserted into *nef*, and encode SIVagm *vif* in the context of the HIV-1 backbone.

Evolutionary Analysis of OWM *A3G*

A3G and mtDNA sequences were subjected to phylogenetic analysis using a Bayesian Monte Carlo Markov Chain (MCMC) approach implemented in BEAST v1.6.1. Sequence alignments were executed using ClustalW. Partial *Cytochrome B* (273 bp) and *NADH1* (298 bp) sequences were concatenated and subjected to 50,000,000 MCMC generations using the HKY85 substitution model, Gamma site heterogeneity model, and estimated base frequencies. The following AGM mtDNA GenBank submissions were included to guide classification efforts: EF597500, EF597501, EF597502, EF597503, AY863426, DQ069713. We used the constant population size coalescent as the tree prior.

The resulting phylogenies were annotated using TreeAnnotator, with a “burnin” of 40 and a posterior probability limit of 0.5. *A3G* sequences were treated similarly, except 100,000,000 MCMC generations were performed. Additionally, phylogenetic analysis by Maximum Likelihood (ML) was performed using the Phylogeny.fr web service. ML trees were imported into PAUP* and trees were scored for log likelihood using the Shimodaira-Hasegawa (SH) test. Test parameters were estimated using 1000 RELL replicates. Alignment of newly derived *A3G* nucleotide sequences from OWM plus those previously published in NCBI GenBank was executed in ClustalW. Phylogenetic reconstruction by maximum likelihood was performed with the web-based version of PhyML [101] (Figure 18). The resulting *A3G* phylogeny and the currently accepted phylogeny of OWM species [17] are similar but not identical, most likely due to rampant selection. The *A3G* phylogeny does not recapitulate that macaques, baboons, and mangabeys share a single common ancestor, and some intraspecies variants do not share immediate common ancestry (De Brazza’s and red capped mangabey). Moreover, Allen’s swamp monkey is placed ancestral to the Cercopithicini tribe (AGM, Patas, guenons), reflecting a previous classification [102]. A phylogeny consistent with the currently accepted phylogeny of OWM species (the placement of Allen’s swamp monkey is the exception) (Figure 18) was uploaded to the Codeml program (of the PAML suite) and to the web-based version of HyPhy (DataMonkey, www.datamonkey.org) for molecular evolution analysis. Given an alignment and phylogenetic tree of primate *A3G*, these packages assess whether or not models of neutral evolution can recapitulate the observed molecular data. *A3G* sequences were screened for recombination with GARD and SBR programs in DataMonkey and the data set was partitioned according to breakpoints [103].

Analyses were performed for the full-length alignment as well as for each partition to consider possible effects of recombination. The mixed effects model of evolution (MEME) analysis was performed in DataMonkey to identify individual codons subject to diversifying selection with a p-value threshold of 0.05, as determined by a significant proportion of branches in the tree exhibiting a bias towards non-synonymous variation at these sites [104]. The MEME analysis is recommended for analyses of diversifying selection in host genes because it is sensitive to cases of transient or episodic selection, whereas traditional methods are not [104]. The Codeml program was used to determine whether *A3G* is evolving under positive selection (comparison of models M7 and M8) and to identify the individual residues undergoing selection (Nsites) [105]. Maximum likelihood scores were calculated under each model and significant differences were calculated using the Chi-square test. Bayes Empirical Bayes (BEB) analysis was used to pinpoint residues with a posterior probability > 0.95 that dN/dS > 1.

Single-Round Viral Infectivity Assays and Western Blot Analysis

293T cells were plated in 12-well plates at 2.5 X 10^5 cells/mL. The following day, cells were cotransfected with 0.4 µg of A3G expression plasmid of an empty expression plasmid, 0.1 µg of L-VSV-G (vesicular stomatitis virus glycoprotein, for pseudotyping), and 0.6 µg of proviral plasmid in a 100 µL transfection volume with TransIT-LT1 lipid transfection reagent (Mirus Bio). Virus supernatants were harvested at 48 hrs and clarified by centrifugation for 5 min at 1,800 rpm, while transfected cells were lysed with NP-40-doc buffer (1% NP-40, 0.2% sodium deoxycholate, 0.12M NaCl, 20 mM Tris [pH 8.0], 2.4 mM dithiothreitol (DTT) and protease inhibitor cocktail (Roche)) and pelleted for 5 min at 10,000 rpm. Total protein concentration was quantified by Bradford assay and 20 µg was resolved by 10% SDS-PAGE, transferred to polyvinylidene difluoride (PVDF) membranes, and probed with anti-HA (Santa Cruz Biotechnology) or anti-actin (Sigma) antibodies. Virus in the supernatant was quantified by p24 Gag enzyme-linked immunosorbent assay (Advanced Bioscience Laboratories). Two ng of virus was used to infect supT1 cells plated at 3.8 x 10^5 cells/mL in the presence of 20 µg/mL DEAE-Dextran, in a total volume of 100 uL. Virus infections were performed in triplicate for 48 hrs. Luciferase activity was measured with 100 µL of Bright-Glo Luciferase Assay Reagent (Promega).

Synthesis of Chimeric A3G and Mutant A3G Expression Plasmids

Chimeric A3G plasmids Chi A, B, and D were produced between Mantled Colobus A3G and AGM haplotype I A3G [98] by restriction digest with BamH1, BstX1, and Apa1, respectively. The remaining chimeras were produced by overlap PCR with reaction-specific primer sets. Mutagenesis of Colobus A3G and chimeras Chi A and C was performed using the Quikchange II XL Site-Directed Mutagenesis Kit (Agilent Technologies).

Accession Numbers

The Genbank accession numbers for Chapter III are as follows: SIVagm Ver90 *vif* sequences isolated from experimentally infected monkeys are JN662632-JN662672. AGM *A3G* sequences are JN662544-JN662631. Accession numbers for OWM *A3G* sequences described in Chapter IV are KC176173-KC176194.

Chapter III: The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict

Introduction

HIV-1 and HIV-2 originated from multiple cross-species transmissions of SIVs that naturally infect nonhuman primates in Africa [106]. Although many natural SIV infections exhibit species-specificity, suggestive of coevolution between virus and host, new viral lineages arise through cross-species transmission and subsequent rounds of virus-host adaptation [45,107]. The adaptation of primate lentiviruses to new hosts is influenced by cell-intrinsic host restriction factors that impose species-specific barriers to viral replication [65].

One such restriction factor is A3G, a cytidine deaminase that inhibits a broad range of retroviruses and retrotransposons [82]. Of the seven members of the APOBEC3 family in humans (A3A-A3H), A3G exhibits the most robust antiviral activity against HIV-1 and is actively expressed in relevant host cells targeted by the virus [108-110]. On the other hand, retroviruses have evolved various mechanisms to overcome the antiviral activity of A3G; for example, most lentiviruses encode the accessory protein Vif*,* which carries out the accelerated degradation of A3G via the proteasome by linking it to a cellular E3 ubiquitin ligase complex [111]. Antagonism of A3G by Vif can occur in a species-specific manner, suggesting that the specificity of Vif reflects adaptation to the host [92,112,113].

The natural SIV infections of AGM provide a unique opportunity to assess how host antiviral genes impact the evolution of lentiviruses and vice versa. AGM are the most abundant non-human primates in Africa, occupying ranges spanning most of the continent south of the Sahara [114]. The AGM species (*Chlorocebus aethiops*) is subdivided into four subspecies commonly known as sabaeus, vervet, tantalus, and grivet monkeys. Despite sharing a recent common ancestor 1-3 million years ago [17,39], each subspecies is naturally infected with a distinct subtype of SIV, aptly named after the specific host it infects: SIVagm.Sab, SIVagm.Ver, SIVagm.Tan, and SIVagm.Gri, respectively [36]. SIVagm infection is widely believed to result in an asymptomatic, non-pathogenic chronic infection in its natural host [46]. Therefore, the host-virus associations in these primates provide a setting to study how viral subtypes have diverged over time as a result of host-specific adaptation in the absence of severe pathogenesis.

In this study, we examine the co-evolution of A3G-Vif interactions. We identify adaptive Single Nucleotide Polymorphisms (SNPs) in *A3G* that allow the antiviral factor to evade destruction by SIVagm Vif, which in turn has triggered *Vif* counter-evolution in a subspecies-specific manner. These data argue that SIV infection selects for Vif-resistant forms of A3G in AGM. Furthermore, in an *in vivo* experimental evolution study, we show that SIVagm Vif proteins evolve adapted specificity that reflects the *A3G* genotype of the host population, implicating the viral gene as a key accessory to the process of host-specific viral adaptation. Our data highlight the conflict-driven dynamics of host-virus arms races and prompts a reexamination of how primate populations are impacted by natural, “non-pathogenic” lentivirus infections.

Results

APOBEC3G is highly polymorphic among AGM subspecies

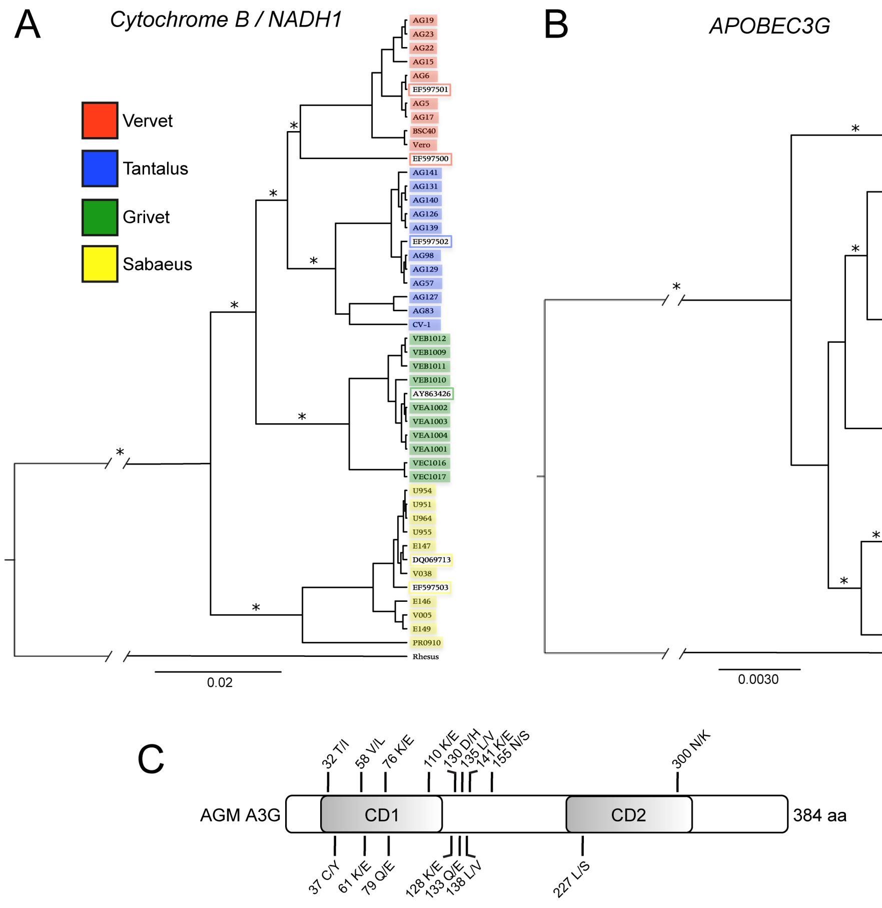
Since *Vif* sequences are among the most divergent viral genes across all lineages of SIV [36], we hypothesized that *Vif* divergence takes place as an adaptive response to differences in the *A3G* repertoire of primate hosts. To determine the degree of *A3G* diversity in different AGM subspecies, we sequenced the gene from 40 unique AGM individuals representing the four major subspecies (Table 1).

ID	Subspecies	Common Name	Origin	SIV Status	Sex	Source	ID	Subspecies	Common Name	Origin	SIV Status	Sex	Source
AG5	*C. pygerythrus*	Vervet	Tanzania	+*	F	VH	AG57	*C. tantalus*	Tantalus	CAR	+	F	MMT
AG6	*C. pygerythrus*	Vervet	Tanzania	+*	F	VH	AG83	*C. tantalus*	Tantalus	CAR	+	F	MMT
AG15	*C. pygerythrus*	Vervet	Tanzania	+*	F	VH	AG98	*C. tantalus*	Tantalus	CAR	+	F	MMT
AG17	*C. pygerythrus*	Vervet	Tanzania	+*	F	VH	AG126	*C. tantalus*	Tantalus	CAR	+	F	MMT
AG19	*C. pygerythrus*	Vervet	Tanzania	+*	F	VH	AG127	*C. tantalus*	Tantalus	CAR	+	F	MMT
AG22	*C. pygerythrus*	Vervet	Tanzania	+*	F	VH	AG129	*C. tantalus*	Tantalus	CAR	+	F	MMT
AG23	*C. pygerythrus*	Vervet	Tanzania	−	F	NPRR	AG131	*C. tantalus*	Tantalus	CAR	−	M	MMT
BSC40	*C. pygerythrus*	Vervet	Cell line	ND	F	ATCC	AG139	*C. tantalus*	Tantalus	CAR	−	F	MMT
Vero	*C. pygerythrus*	Vervet	Cell line	ND	F	ATCC	AG140	*C. tantalus*	Tantalus	CAR	−	F	MMT
E146	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	M	CA	AG141	*C. tantalus*	Tantalus	CAR	+	F	MMT
E147	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	F	CA	CV-1	*C. tantalus*	Tantalus	Cell line	ND	M	ATCC
E149	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	F	CA	VEA1001	*C. aethiops*	Grivet	Ethiopia	ND	F	NF
V038	*C. sabaeus*	Sabaeus	St. Kitt’s	−	M	NPRR	VEA1002	*C. aethiops*	Grivet	Ethiopia	ND	F	NF
V005	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	F	VH	VEA1003	*C. aethiops*	Grivet	Ethiopia	ND	F	NF
U951	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	F	VH	VEA1004	*C. aethiops*	Grivet	Ethiopia	ND	F	NF
U954	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	F	VH	VEB1009	*C. aethiops*	Grivet	Ethiopia	ND	F	NF
U955	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	F	VH	VEB1010	*C. aethiops*	Grivet	Ethiopia	ND	M	NF
U964	*C. sabaeus*	Sabaeus	St. Kitt’s	+*	F	VH	VEB1011	*C. aethiops*	Grivet	Ethiopia	ND	M	NF
PR09010	*C. sabaeus*	Sabaeus	St. Kitt’s	ND	M	Coriell	VEB1012	*C. aethiops*	Grivet	Ethiopia	ND	F	NF
							VEC1016	*C. aethiops*	Grivet	Ethiopia	ND	F	NF
							VEC1017	*C. aethiops*	Grivet	Ethiopia	ND	M	NF

Table 1. African Green Monkey sample cohort
A total of 40 AGM-derived nucleic acid samples were obtained from PBMC or continuous cell lines. SIV status indicates evidence of natural infection (+) or experimental infection in previous study (+*). For some monkeys, SIV status has not been determined (ND). Samples supplied by: VH, Vanessa Hirsch; CA, Cristian Apetrei; MMT, Michaela Muller-Trutwin; NF, Nelson Freimer. The origin of these samples is as follows: AG23 and V038 samples were provided by the Nonhuman Primate Research Resource (NPRR). PR0911 was obtained from Coriell Biorepository. The CV-1, BSC40, and Vero cell lines were obtained from ATCC. Vervet monkeys were wild caught in Tanzania and transferred to an NIH animal facility [115,116]. Tantalus monkeys were wild caught in the Central African Republic (CAR) and transferred to the Institut Pasteur in Bangui [117]. Caribbean AGM with ancestry tying back to western Africa were wild caught on the island of St. Kitt’s and transferred to the Tulane National Primate Research Center. These sabaeus monkeys were described in [118] with the support of R01 AI064066 to Ivona Pandrea. Grivet monkeys were wild caught in Ethiopia and sedated for sampling before being returned to their habitat. The samples are derived from the Vervet Genome Project with the support of R01 RR016300 to Nelson Freimer.

Mitochondrial genes (mtDNA) were also sequenced in order to confirm the subspecies of origin of each AGM sample. AGM mtDNA sequences exhibited monophyly; sequences derived from the same AGM population clustered together to form clades related by common ancestry (Figure 10a). In contrast, *A3G* sequences from any given AGM subspecies are paraphyletic (Figure 10b), indicating that the evolutionary history of *A3G* is incongruent with the species phylogeny.

Virtually every polymorphism in AGM *A3G* yields an amino acid replacement—we identified 28 non-synonymous SNPs (nsSNPs) present in more than one individual, 16 of which were observed with a minor allele frequency of 5% or greater (Figure 10c). In contrast, only one prevalent synonymous SNP was found, and it was tightly linked with a neighboring nsSNP. The heavy bias of non-synonymous changes in AGM *A3G* is suggestive of adaptive evolution by natural selection. Therefore, to address the consequences of AGM *A3G* polymorphism, we selected 2 haplotypes from each of the four subspecies for functional analysis (Figure 10b and 10c)**.** *A3G* haplotypes were co-transfected with proviral lentiviruses in single-round infectivity experiments to measure their inherent antiviral activity. In this assay, the infectivity of each virus produced in the absence of A3G is normalized to 100%. Hereafter, the proteins encoded by *A3G* haplotypes are referred to as A3G variants.



DNA	95	110	172	181	226	235	328	382	388	397	403	412	421	465	680	900
Amino acid	32	37	58	61	76	79	110	128	130	133	135	138	141	155	227	300
Hap I	I	C	V	K	K	Q	E	K	D	Q	L	L	E	S	S	N
Hap II	T	Y	L	E	K	Q	K	K	D	Q	L	L	E	N	S	N
Hap III	T	Y	L	E	E	Q	K	K	D	Q	L	L	K	N	L	N
Hap IV	T	Y	L	K	K	Q	E	K	D	Q	L	L	K	N	L	N
Hap V	I	C	L	E	K	E	E	K	D	Q	L	L	K	N	L	N
Hap VI	I	C	L	E	K	Q	E	K	D	E	V	V	K	N	L	K
Hap VII	I	C	L	E	K	Q	K	E	D	Q	L	L	K	N	S	N
Hap VIII	I	C	L	E	K	Q	E	K	H	E	L	L	K	N	S	N

Figure 10. Phylogenetic analysis of partial mitochondrial genomes and *APOBEC3G*
The evolutionary relationship between AGM samples was inferred by Bayesian MCMC phylogenetic reconstruction as described in the Methods. Sequence IDs are color-coded according to legend to indicate AGM subspecies of origin. Rhesus macaque DNA was used as an outgroup. Ancestral outgroup branch length is shortened due to space constraint. Scale bars indicate an arbitrary unit of time, relative to 1. (A) Partial mitochondrial DNA sequences. Asterisks (*) mark nodes with posterior probability scores of ≥0.9. (B) Full-length *A3G* sequences from AGM. Asterisks (*) mark nodes with posterior probability scores of ≥0.5. Eight distinct *A3G* haplotypes selected for functional analysis are indicated by I-VIII. (C) Sixteen nsSNPs observed at a frequency ≥5% are indicated according to amino acid position, relative to the N-terminal cytidine deaminase domain (CD1) and the C-terminal cytidine deaminase domain (CD2). The eight AGM *A3G* haplotypes selected for functional analysis encode proteins that are differentiated by amino acid changes at these positions.

Our results indicate that each A3G variant exhibits strong restriction of SIVagm.TanΔVif and HIV-1ΔVif (Figure 11a), decreasing viral infectivity of HIV-1ΔVif on average by approximately 150-fold. For comparison, HIV-1ΔVif infectivity was inhibited about 140-fold by human A3G. These data suggest that there is extensive polymorphism in AGM *A3G,* and that this diversity does not affect intrinsic antiviral activity against lentiviruses.

Evidence for antagonism-driven evolution of A3G in AGM

In order to address whether the polymorphism in AGM *A3G* affects susceptibility to the viral antagonist Vif, we constructed four recombinant HIV-1 proviruses expressing Vif from the following SIVagm subtype viruses: SIVsab-1 [37], SIVgri.667 [119], SIVver90 [120], and SIVtan.1 [121]. *A3G* haplotypes were co-transfected with HIV-1 provirus expressing SIVagm Vif and virus infectivity was assayed in single-cycle infections. The antiviral activity of A3G variants I-VI is readily counteracted by viruses encoding each SIVagm Vif protein, resulting in a rescue of viral infectivity that approaches levels observed in the absence of A3G (Figure 11b). This finding suggests that the SNPs contained in haplotypes I-VI do not affect sensitivity to Vif.

Haplotypes VII and VIII are unique in that they encode single amino acid changes at codons 128 and 130, respectively (Figure 3A), which in the context of the motif ^{126}FWKPDYQ^{132} have been implicated in the interaction between human A3G and HIV-1 Vif [95,122]. We found that variant VII is resistant to Vif from SIVagm.Ver, SIVagm.Tan, and SIVagm.Sab (sensitive only to SIVagm.Gri Vif), while variant VIII is also resistant to Vif from SIVagm.Ver and SIVagm.Tan (Figure 3B). Interestingly, haplotypes VII and VIII are confined to single subspecies of AGMs (the grivet and sabaeus monkeys, respectively) at an allele frequency of 20% (4/20 chromosomes per subspecies) (Figure 3A).

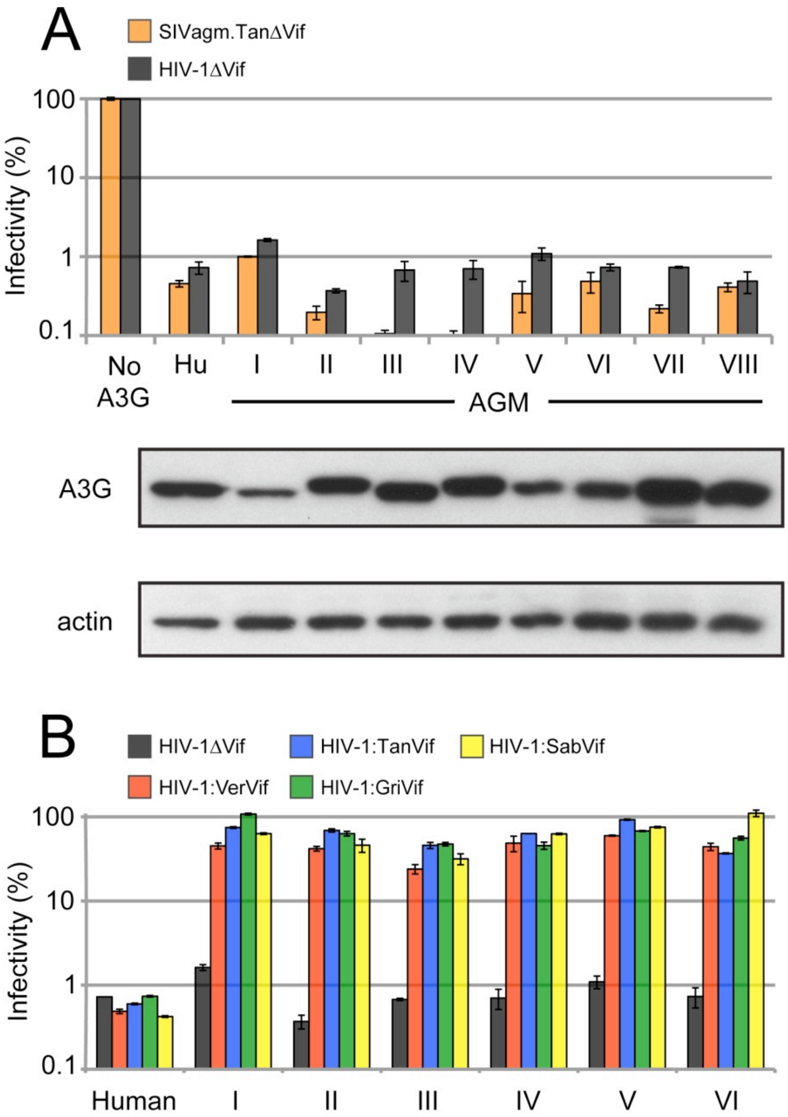


Figure 11. AGM A3G variants maintain potent restriction of lentiviruses
Single-round viral infectivity assays were performed with virus produced in the presence or absence of A3G proteins. Error bars indicate standard deviation from the mean of three infection replicates. (A) Infectivity of Vif-deficient SIVagm.Tan and HIV-1 (SIVagm.TanΔVif and HIV-1ΔVif). Infectivity in the absence of A3G was normalized to 100% (No A3G), while infectivity in the presence of human A3G served as a positive control for virus restriction. Anti-HA Western blot analysis was used to measure expression of AGM A3G variants in 293T cells. Anti-β-actin served as protein loading controls. (B) Infectivity of HIV-1ΔVif and HIV-1 expressing SIVagm Vif in the presence of AGM A3G variants I-V.

To demonstrate that the Vif-resistance exhibited by these two variants of A3G (henceforth referred to as variants VII K128E and VIII D130H) (Figure 12a) is conferred by single residue changes at sites 128 and 130, respectively, we performed site-directed mutagenesis to revert the amino acid at these sites to the more common residues present in variants I-VI. An E128K change rendered variant VII K128E fully susceptible to all four SIVagm Vif proteins, confirming the importance of this site for Vif recognition (Figure 12b). Similarly, a H130D mutation rendered variant VIII D130H fully sensitive to SIVagm Vif (Figure12b). Western blot analysis shows that antagonism of A3G variants by Vif results in depletion of intracellular A3G protein. A3G variants VII K128E and VIII D130H are effectively degraded by SIVagm.Gri Vif and SIVagm.Sab Vif, respectively, but not by SIVagm.Ver Vif (Figure 12c). Furthermore, the revertant mutations E128K and H130D confer sensitivity to degradation by SIVagm.Ver Vif, mirroring the results of the infectivity rescue experiments (Figure 12c). Thus, standing variation in the *A3G* locus of AGM is adaptive, as it allows the antiviral factor to evade destruction by the viral antagonist, Vif. Furthermore, the different specificities exhibited by SIVagm Vif proteins demonstrate that SIVagm lineages have co-adapted to distinct versions of A3G in AGM subspecies.

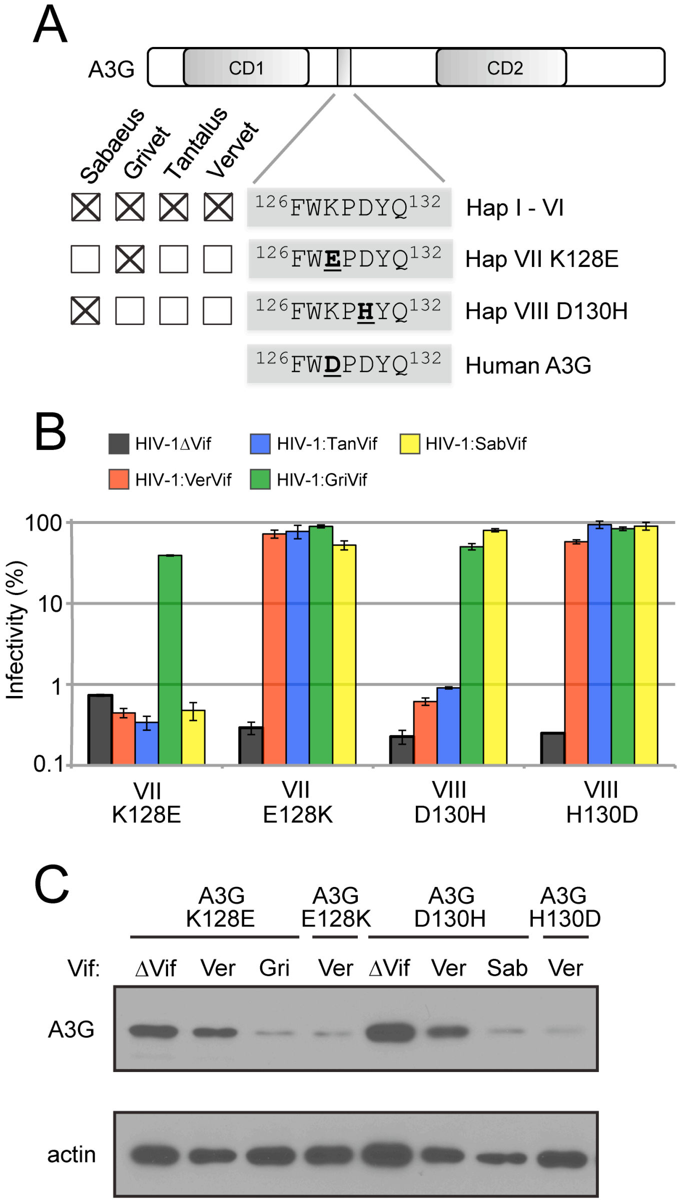


Figure 12. Vif-resistance phenotype of AGM A3G maps to residues 128 and 130 of the putative Vif binding site

(A) The presence of *A3G* haplotypes in AGM subspecies is indicated by checked boxes. AGM *A3G* haplotypes VII and VIII, exclusive to grivet and sabaeus monkeys, respectively, are distinguished by a negatively charged glutamic acid at residue 128, and a positively charged histidine at residue 130, respectively. Human A3G encodes a negatively charged aspartic acid at residue 128.

(B) Single-round viral infectivity assays were performed with HIV-1ΔVif and HIV-1 expressing SIVagm Vif in the presence of AGM A3G variants VII K128E and VIII D130H. Error bars indicate standard deviation from the mean of three infection replicates.

Mutagenesis of sites 128 and 130 restored Vif sensitivity to Vif-resistant AGM A3G variants VII K128E and VIII D130H, respectively.

(C) Anti-HA Western blot analysis of A3G from cell lysates co-transfected with HIV-1 encoding different Vif proteins. Anti-β-actin served as protein loading controls.

In humans, residue 128 of A3G is occupied by an aspartic acid residue that is presumably fixed in global populations [123,124] (Figure 13). By virtue of possessing an acidic residue at this site, we suspected that A3G variant VII K128E would be susceptible to antagonism by Vif from human lentiviruses, namely HIV-1 Vif. By co-transfecting agmA3G variants with HIV-1 proviral plasmid, we show that variant VII K128E is uniquely susceptible to HIV-1 Vif, reinforcing the importance of residue 128 in the specificity with which Vif antagonizes A3G (Figure 13). We have thus identified the first naturally occurring form of A3G in Old World Monkeys (OWM) that is sensitive to HIV-1 Vif. Further dissection of the host restriction factor repertoire encoded by these animals will determine if they represent novel candidates in the search for a non-human primate model of HIV-1 infection.

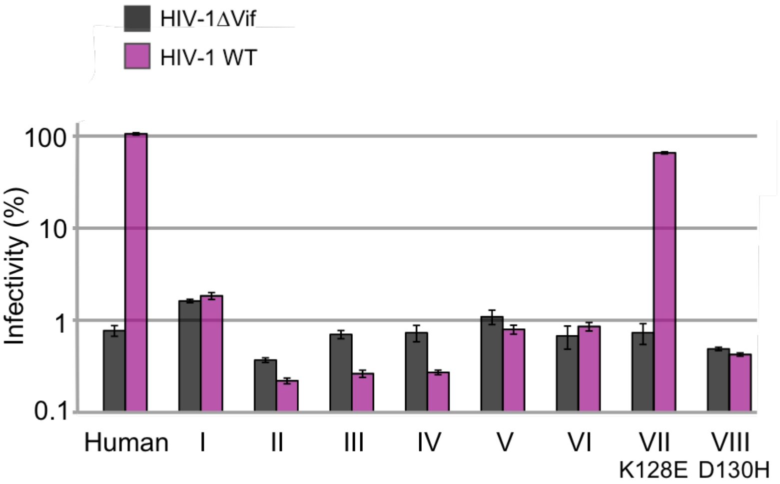


Figure 13. An A3G variant specific to grivet monkeys is sensitive to HIV-1 Vif
Single-round viral infectivity assays were performed with virus produced in the presence or absence of A3G proteins. Error bars indicate standard deviation from the mean of three infection replicates. Infectivity of Vif-deficient HIV-1 and HIV-1 expressing HIV-1 Vif in the presence of AGM A3G variants I-VIII. Infectivity in the absence of A3G was normalized to 100% (No A3G), while infectivity in the presence of human A3G served as a positive control for HIV-1 Vif activity.

Interestingly, our data shows that while HIV-1 Vif is capable of antagonizing variant VII K128E, SIVagm.Gri Vif was incapable of antagonizing human A3G (Figure 11b). This demonstrates that, although residue 128 is a critical mediator of the A3G-Vif interaction, it is not the sole determinant. Therefore, the full spectrum of Vif-interacting sites on A3G remains to be elucidated.

Experimental adaptation of SIVagm to Vif-resistant forms of AGM A3G in vivo

Retrospective studies of virus evolution allow one to infer adaptations to a host, but do not address the timeframe in which this occurs. On the other hand, prospective studies in which the time of infection is known can be powerful tools in understanding the dynamics of viral adaptation. We took advantage of an experimental evolution study in which the impact of *A3G* variation on SIVagm evolution could be evaluated within the timeframe of single infected animals. In a previous study, sabaeus monkeys were experimentally infected with SIVagm.Ver90 and plasma viremia was monitored in infected sabaeus monkeys for nearly two years [115]. We genotyped the *A3G* locus in these animals and found that monkey V005 is homozygous for *A3G* haplotype VIII D130H, which is resistant to SIVagm.Ver90 Vif (Figure 12b), while monkey V038 is heterozygous for D130H. Moreover, monkey U964 is homozygous for A3G D130, which is found in A3G variants that are sensitive to SIVagm.Ver90 Vif (Figure 11b**).** As this Vif is derived from the same virus used to inoculate the monkeys experimentally, we anticipated a scenario in which restriction by a Vif-resistant form of A3G would select for a functional form of Vif *in vivo.* Thus, we had the opportunity to track *vif* evolution in monkeys expressing different A3G variants.

We amplified *vif* from plasma viral RNA isolated 6-12 months post-infection. Phylogenetic analysis of >10 viral sequences per monkey demonstrates that *vif* isolates sort by *A3G* genotype, with *vif* from monkeys V005 (H130/H130) and U964 (D130/D130) exhibiting monophyly and *vif* from V038 (D130/H130) forming two distinct clades (Figure 14a). An alternate topology in which V038 *vif* share a single ancestor was rejected by the SH-test ($p = 0.006$) [125]. The SIVagm.Ver90 *vif* sequence lies near the hypothetical ancestor generated by the analysis. Three to five unique *vif* sequences that were representative of the diversity seen within each monkey were cloned into an HIV-1 backbone and their ability to antagonize A3G variant VIII D130H in single-round infectivity assays was measured (Figure 14a, black stars). All three Vif isolates derived from an infected monkey expressing two copies of A3G variant VIII D130H (V005) exhibited robust antagonism of the this variant, conferring a 38-fold increase in viral infectivity on average relative to the unadapted control SIVagm.Ver90 Vif ($p < 0.001$)(Figure 14b, left). By comparison, Vif isolated from monkeys expressing one or zero copies of the haplotype VIII D130H (monkeys V038 and U964, respectively) did not demonstrate elevated antagonism of this A3G variant relative to SIVagm.Ver90 Vif ($p > 0.05$). We found that all of the *in vivo* Vif isolates antagonized the AGM *A3G* haplotype VI encoding the ancestral D130 allele, with the exception of some *vif* genes derived from monkey V005 (Figure 14b, right). Together, these results demonstrate that the *vif* sequences tested encode functional proteins, and moreover, that observed differences in antagonism are due to changes that alter specificity for the A3G substrate.

We also expressed Vif isolates in the presence of different amounts of A3G variant VIII D130H. Again, we observed that Vif from a monkey homozygous for this haplotype (V005) exhibits robust antagonism at all concentrations of A3G (Figure 14c, white triangles). However, the adaptation is partial, as infectivity is not rescued to the same degree as SIVagm.Sab Vif, the fully adapted control that has undergone repeated transmission cycles within wild sabaeus populations (Figure 14b and 14c, black squares).

Conversely, Vif from the heterozygous monkey (V038) and the monkey homozygous for D130 (U964) do not antagonize variant VIII D130H at any concentrations relative to SIVagm.Ver90 Vif (Figure 14c, white squares and white circles versus black triangles). Consistent with the infectivity data, V005 Vif effectively degrades A3G variant VIII D130H and excludes it from virions, while Vif from U964 does not (Figure 14d). Therefore, there is evidence of adaptation in the *vif* gene upon infection of a single animal, the extent of which is influenced by the host *A3G* genotype.

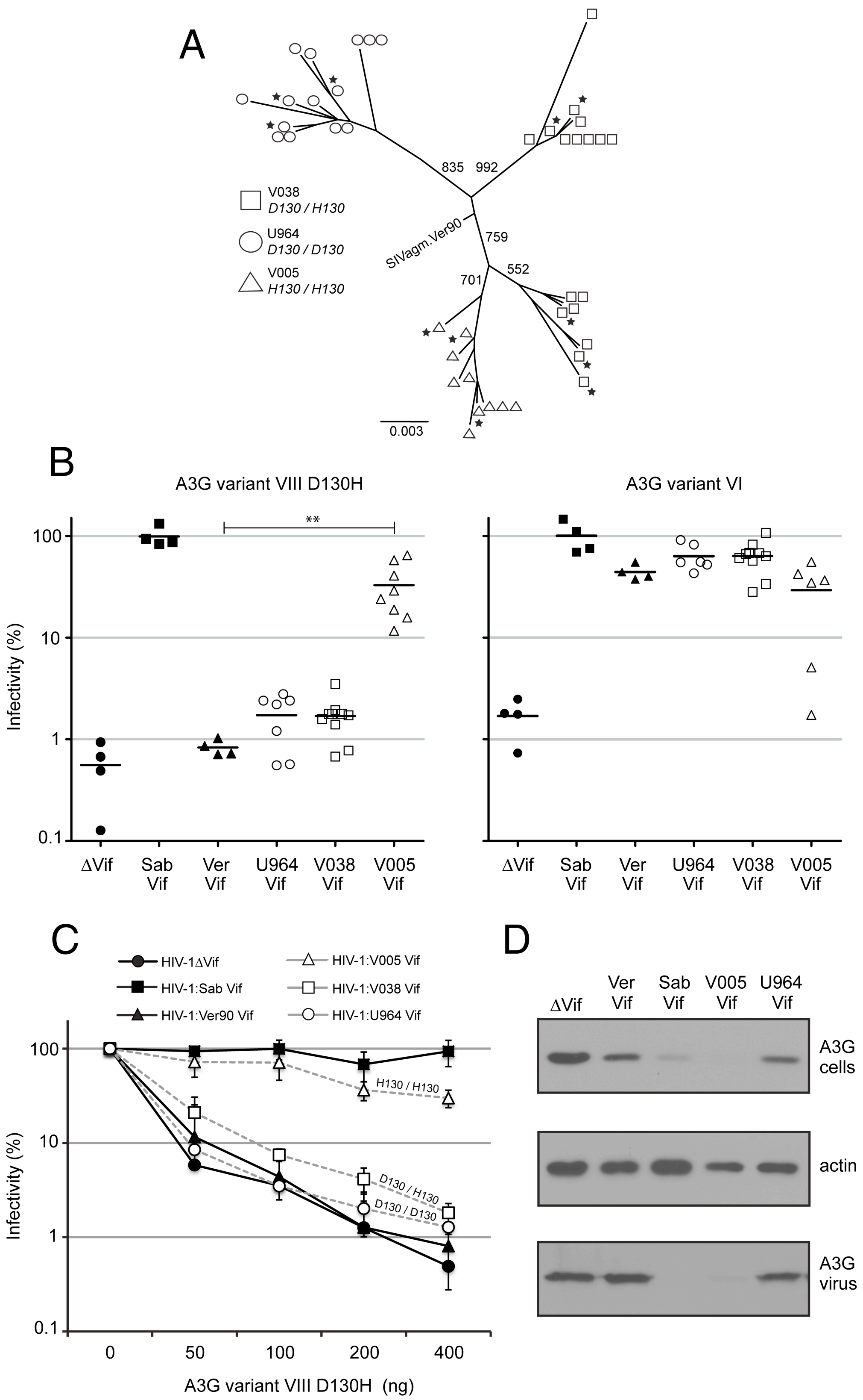

A
V038
D130 / H130
U964
D130 / D130
V005
H130 / H130
SIVagm.Ver90
835
992
759
552
701
0.003
B
A3G variant VIII D130H
A3G variant VI
Infectivity (%)
100
10
1
0.1
**
ΔVif
Sab Vif
Ver Vif
U964 Vif
V038 Vif
V005 Vif
C
HIV-1ΔVif
HIV-1:Sab Vif
HIV-1:Ver90 Vif
HIV-1:V005 Vif
HIV-1:V038 Vif
HIV-1:U964 Vif
H130 / H130
D130 / H130
D130 / D130
Infectivity (%)
0
50
100
200
400
A3G variant VIII D130H (ng)
D
ΔVif
Ver Vif
Sab Vif
V005 Vif
U964 Vif
A3G cells
actin
A3G virus

Figure 14. Experimental adaptation of SIVagm to Vif-resistant forms of AGM A3G *in vivo*
(A) *Vif* sequences amplified by nested RT-PCR were aligned and subjected to phylogenetic analysis. An unrooted, neighbor-joining plot was generated after 1000 iterations, with bootstrap values indicated above branches. The scale bar indicates the number of nucleotide substitutions per site. *Vif* sequences were isolated from two time points post-infection: 6 months (circles) and 12-months (triangles). The sequence for SIVagm.Ver90 *vif* is included to show proximity to the hypothetical ancestor. Blacks stars indicate sequences cloned into the HIV-1ΔVif background and tested for activity against A3G. Three to five unique *vif* sequences amplified from each of three sabaeus monkeys were cloned into the HIV-1ΔVif background. HIV-1ΔVif is denoted by black circles. See also Figure for Vif alignments. (B) Left: Single-round viral infectivity assays were performed with virus produced by co-expressing these HIV-1:SIV Vif proviral plasmids with AGM A3G variant VIII D130H. The plotted points per column correspond to infectivity experiment replicates performed with viruses encoding Vif isolated from monkey plasma (each virus replicated two to four times) except in the case of No Vif, SIVagm.Sab Vif, and SIVagm.Ver Vif, in which plotted points represent three independent experiments using the same virus. Vif from SIVagm.Sab is denoted by black squares; SIVagm.Ver, black triangles; V005 (two copies of D130H), white triangles; V038 (one copy of D130H), white squares; U964 (zero copies of D130H), white circles. Horizontal bars indicate the mean relative infectivity. *p*-values were calculated using a one-way ANOVA test, relative to SIVagm.Ver90 Vif. Right: Single-round viral infectivity assays were performed with virus produced by co-expressing the same HIV-1:SIV Vif proviral plasmids with Vif-sensitive AGM A3G variant VI, encoding D130. (C) Single-round viral infectivity assays were performed with virus produced by co-expressing HIV-1ΔVif or HIV-1:SIV Vif proviral plasmids in the presence of decreasing concentrations of AGM A3G variant VIII D130H. Each point corresponds to the mean infectivity of 3-5 viruses encoding different Vif proteins, each replicated 2-4 times each, except in the case of No Vif, SIVagm.Sab Vif and SIVagm.Ver Vif. (D) Western blot analysis was used to measure A3G protein expression in cells (upper panel) and encapsidated in virions (lower panel) after co-transfection with HIV-1 encoding Vif derived from animals V005 and U964. Anti-β-actin served as cellular protein loading controls.

Amino acid alignment of all Vif isolated from these monkeys reveal a number of recurrent mutations that arose during the time course of infection, and in some cases, the same mutation appeared independently in different animals (Figure 15a). However, only one mutation (Y84C) was observed in 100% (10/10) of Vif isolates from monkey V005 but not in a single Vif obtained from the other two monkeys. Interestingly, the equivalent change is not observed in SIVagm.Sab Vif, demonstrating that Vif may adapt via a number of distinct evolutionary trajectories (Figure 15b).

Figure 15. Amino acid alignment of Vif sequences isolated from infected monkey plasma and the parental virus strain, SIVver90

(A) Vif isolates from monkey V038 (n=17, orange), V005 (n=10, green), and U964 (n=14, blue) were aligned with amino acid differences shown relative to SIVagm.Ver90 Vif. The Y84C mutation associated with antagonism of the AGM A3G variant VIII D130H is outlined in a red box. Blacks stars indicate sequences cloned into the HIV-1 background and tested for activity against A3G.

(B) The amino acid differences between SIVagm.Ver90 Vif and SIVagm.Sab Vif.

Some Vif sequences isolated from monkey V005 demonstrating a gain-of-activity against A3G variant VIII D130H exhibited a coincident loss-of-activity against ancestral A3G, suggesting that *vif* counter-evolution results in a trade-off (Figure 14b and Figure 16, dark gray). To assess the effect of Y84C on the function of Vif, we reverted the mutation back to the ancestral character present in the parental virus strain and measured its activity against ancestral and derived (Vif-resistant) A3G variants. Mutagenesis of position 84 (C84Y) abrogated the ability of Vif to antagonize variant VIII D130H, demonstrating that a single amino acid replacement conferred a gain-of-function against a formerly Vif-resistant form of A3G *in vivo* (Figure 16, light gray). Interestingly, the C84Y revertant conferred Vif with superior activity against an ancestral A3G variant (variant VI) (Figure 16, light gray). This suggests that the Y84C signature emerged *in vivo* to enable antagonism of a novel, derived form of A3G, but at the cost of antagonizing the ancestral form. The virus evolving within monkey V005, however, did not incur this cost because its host does not express the ancestral A3G variant. Therefore, we have identified an evolutionary trade-off that may provide an explanation for why *vif* did not counter-evolve in monkey V038, which is heterozygous for A3G (containing both the ancestral and derived variants).

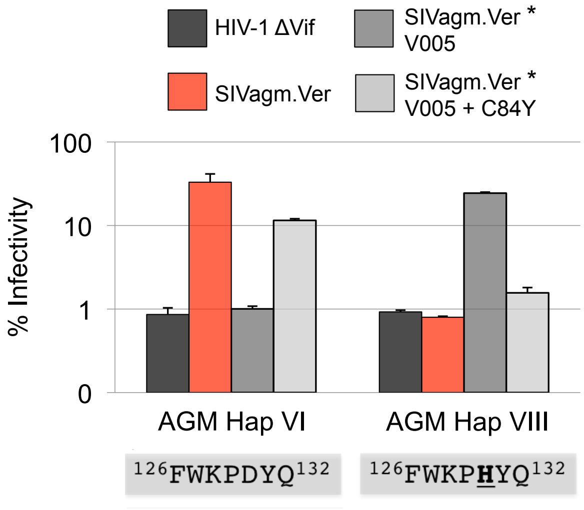


Figure 16. A 'costly' adaptive mutation in SIVagm Vif
The Y84C signature common to Vif isolates that counter-evolved *in vivo* was reverted by site-directed mutagenesis and tested for activity against two variants A3G in the autologous host. SIVagm.Ver represents Vif from the parental strain (ancestral), while SIVagm.Ver* is derived Vif isolated from infected monkey V005.

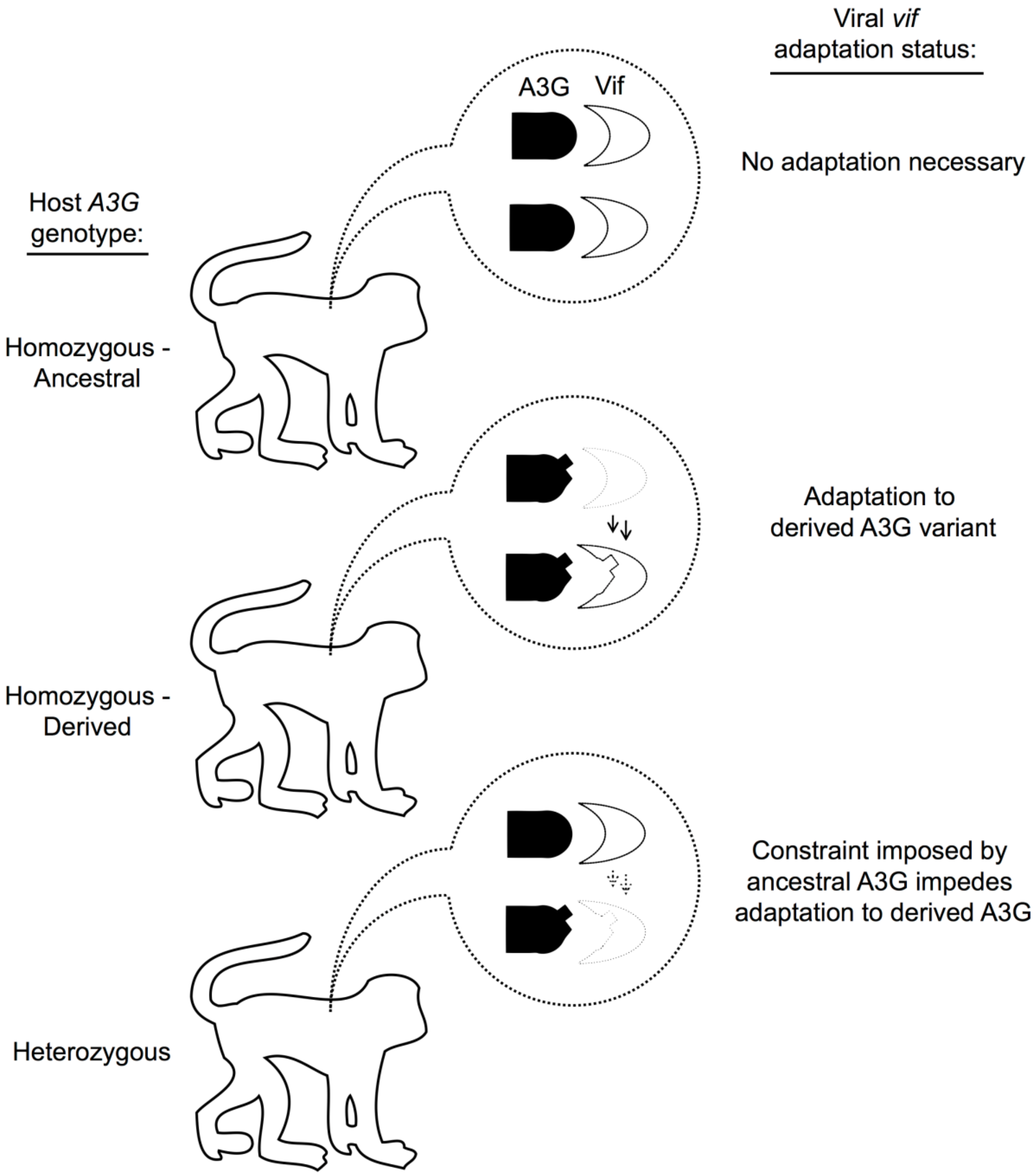


Figure 17. Host heterozygosity may impede viral adaptation in the setting of cross-species transmission.
The outcome of experimental infection of three AGM that differ by *A3G* genotype is schematized. A3G variants are drawn in black, while Vif is drawn in white. Arrows signify *in vivo* evolution of the *vif* gene.

Discussion

We present evidence that *A3G* is adaptively diversifying within AGM as a result of antagonism-driven evolution, suggesting a history of ongoing host-virus conflict. Specifically, we found two sites located in the Vif binding site of A3G that are functionally polymorphic in AGM: high frequency, charge-altering SNPs at residues 128 and 130 alter the sensitivity of A3G to Vif, allowing the antiviral factor to evade antagonism. In response to changes in host A3G, we demonstrate using both natural isolates and experimental evolution *in vivo* that the viral-encoded Vif protein must itself adapt to re-target the restriction factor and carry out its destruction. We predict that the genetic conflict between these two proteins drives continuous evolution in *vif*, which contributes to the species-specificity of natural SIV infections.

Selection costs incurred by apparently "non-pathogenic" viruses

Evolutionary studies indicate that *A3G* is rapidly evolving in different primate species through a process of positive selection [90,126]. The Vif-resistant SNPs K128E and D130H were observed at 20% frequency in *A3G* of the grivet and sabaeus subspecies, respectively, suggesting that they have emerged recently relative to other polymorphisms that were identified in multiple AGM subspecies. Despite the evidence that contemporary infections result in a non-pathogenic outcome, the appearance of adaptive genetic changes in the host implies that SIV infection has exerted selective pressure in at least some AGM populations. It is possible that selection imparted by SIVagm may manifest overtly as progressive disease and immunodeficiency in some monkeys. In support of this, immunodeficiency has been observed in captivity among models of “natural” SIV infection [127-129]. The virus may also affect fitness more subtly by reducing female fertility, influencing mate choice, or increasing infant mortality [130]. Alternatively, SIVagm may have been pathogenic to AGM hosts in the past. Even so, our characterization of the ongoing genetic conflict between SIVagm and AGM suggests that other “non-pathogenic” primate lentiviruses may select for adaptive changes in their respective host species as well.

Paraphyletic phylogeny of A3G in African Green Monkey subspecies and Vif selection

The failure of *A3G* sequences to fully segregate by AGM subspecies may result from demographic and/or stochastic processes. For example, male-biased migration of AGM between neighboring reproductive groups can lead to rampant gene flow in nuclear genes but not in mtDNA, which is maternally-inherited from the more sedentary AGM females [131]. In addition, incomplete lineage sorting could result in alternate retention of ancestral *A3G* alleles, leading to an apparent paraphyly. Alternatively, the incongruent *A3G* phylogeny may be explained by selection. The fact that residues in A3G that interact with Vif are undergoing positive selection in primates [126] is further evidence that Vif-mediated antagonism is capable of selecting for genetic change in *A3G.* In addition, long-term balancing selection can maintain ancestral gene polymorphism across diverged populations due to a survival or reproductive advantage conferred to heterozygous individuals.

The adaptive inferiority of Vif isolated from monkey V038, despite expressing one copy of A3G variant VIII D130H (Figure 14), suggests that *A3G* heterozygosity may apply an adaptive constraint on the virus. The relative lack of viral adaptation in this animal demonstrates that the broadening of Vif specificity in order to accommodate multiple A3G variants may come at a considerable cost, which is likely to decelerate viral adaptation. This is supported by the branching order of the *vif* phylogenetic tree, in which V038 *vif* sequences are split into two clades (Figure 14a, triangles). We postulate that a lack of directional selection in these viral sequences results from the requirement that Vif must counteract two distinct A3G variants in a heterozygous animal.

Host-virus adaptation

A recent report demonstrated that the outcome of experimental cross-species transmission of SIV in rhesus macaques is largely dependent on the TRIM5α genotype of the recipient monkey [71]. Furthermore, TRIM5α-resistant virus emerged in some of the infected monkeys, demonstrating that adaptation in the capsid (CA) region of Gag facilitates escape from TRIM5α-mediated restriction *in vivo*. We employed a similar approach to demonstrate that *vif* counter-evolves *in vivo* to degrade a formerly Vif-resistant variant of A3G. While experimental infections are unlikely to reproduce the natural dynamics of host-virus coevolution, we show that A3G-driven evolution of *vif* is a naturally occurring phenomenon because the specificity of Vif proteins from SIVagm subtypes reflects adaptation to subspecies-specific A3G variants in AGM populations exposed to SIV in the wild rather than in the laboratory.

There is not a direct correlation between *A3G* genotype and viral load in this small cohort [115]. Thus, A3G is not the sole determinant of clinical outcome in this virus-host setting. Nonetheless, by tracking viral evolution in SIV-infected monkeys, we have shown that adaptation in *vif* occurs upon infection of monkeys expressing naturally occurring variants of A3G, demonstrating that A3G has the potential to set the course of viral evolution *in vivo*. By analyzing intraspecies *A3G* variation in a natural host of SIV, we have now demonstrated a case of contemporary selection and adaptation between primate lentiviruses and their natural primate hosts.

Chapter IV: Convergence and divergence in the evolution of the APOBEC3G-Vif interaction reveal ancient origins of Simian Immunodeficiency Viruses

Introduction

Previously, we studied the co-evolution of *A3G* and *vif* in the setting of natural SIVagm infections in AGM. We found that the *A3G* is adaptively diversifying in wild monkey populations, with single nucleotide polymorphisms (SNPs) encoding charge altering amino acid changes at surfaces targeted by Vif [98]. Our data support that these naturally occurring mutations in *A3G* were selected to allow evasion of Vif proteins, implicating SIV Vif or a Vif-like factor as the selective pressure responsible. Adaptive evolution at the A3G-Vif interface implies that SIV infections can incur a cost to host fitness, whether it be overt immunodeficiency or more subtle phenotypes that decrease host survivability or fertility [67,98].

In the present study, we expand our analysis of co-evolution to multiple primate species and multiple SIV lineages in order to estimate the age of the primate-lentivirus relationship. Here, we report that multiple primate species of the Old World Monkey (OWM) subfamily *Cercopithecinae* possess mutations in the Vif-binding site of *A3G* and that each allows escape from antagonism by Vif proteins. The recurrence and deep ancestry of such mutations suggest that a lentivirus encoding Vif existed at least 5-6 MYA. Moreover, we reveal an even older ancestral insertion event in the N-terminus of *A3G* of the *Colobinae* subfamily that conceals the Vif-binding site and precludes interaction with Vif proteins, suggesting that lentiviruses may have infected primates as much as 12 MYA. In response, contemporary Vif proteins have counter-evolved to these various Vif-resistant forms of A3G by either toleration of amino acid variation at the canonical Vif-binding site or by engagement of non-canonical surfaces of A3G. Furthermore, some Vif proteins have evolved broad cross-reactivity during their long-time adaptation to primate hosts. Vif from a lentivirus infecting sooty mangabeys (SIVsm) and its descendants (SIVmac and HIV-2), exhibit exceptional breadth, possibly explaining in part how SIVsm was able to successfully colonize both humans and macaque species. Together, these data suggest that primate lentivirus infections are older than previous thought, driving selective changes in antiviral genes of their natural hosts and inciting an evolutionary arms race that continues to this day.

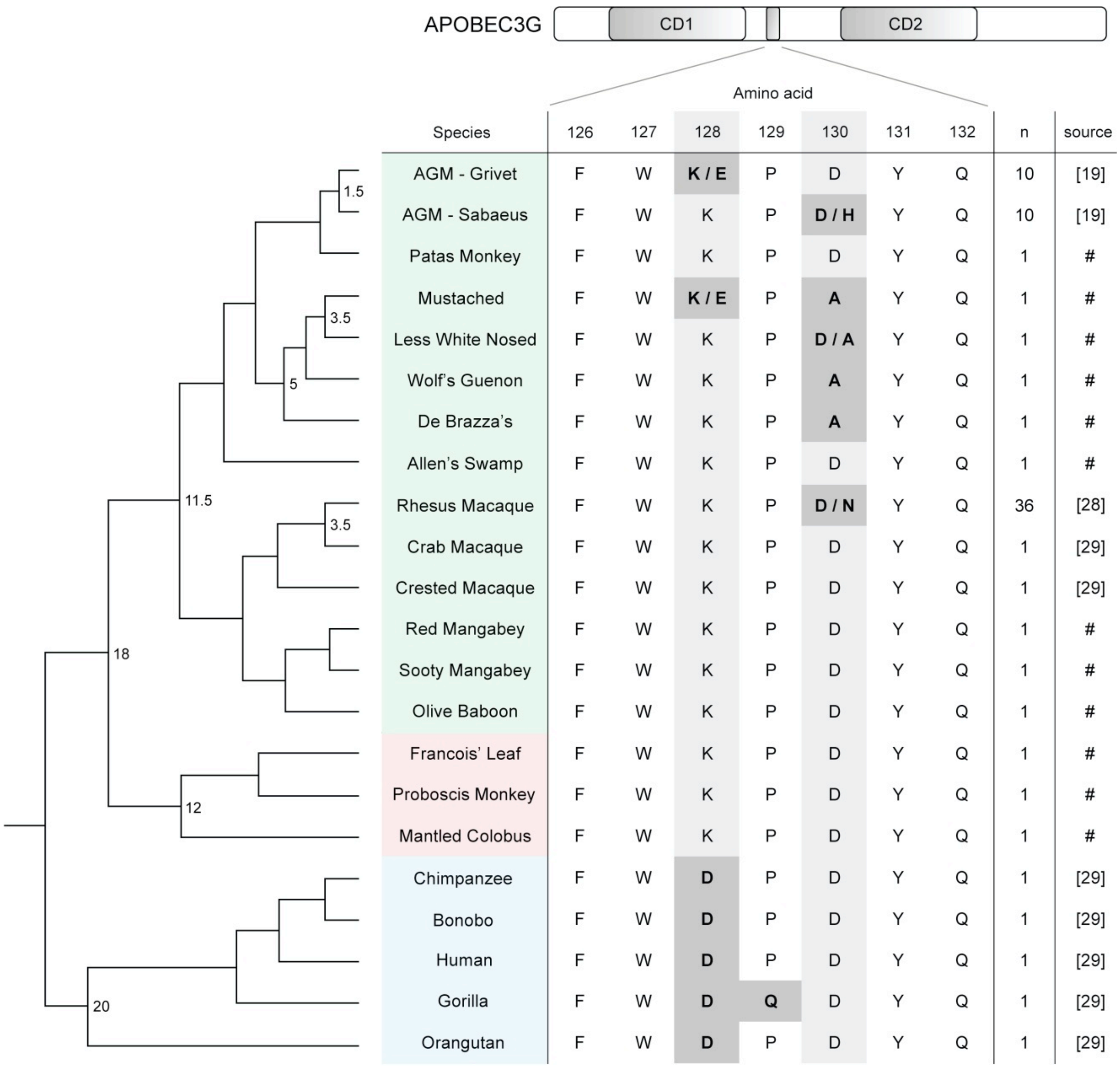
Results

The Vif binding site of A3G is adaptively diversifying in the Cercopithecinae subfamily of Old World Monkeys

The bulk of known SIVs that circulate in the wild have been found within the *Cercopithecinae* subfamily of OWM, a group that includes AGM, mangabeys, macaques, and members of the *Cercopithecus* genus (collectively known as guenons) [36]. We previously identified naturally occurring amino acid changes within ^{126}FWKPDYQ^{132} of

A3G in AGM populations, a motif that is critical to the interaction between A3G and Vif [93,113,132]. Specifically, polymorphisms at codons 128 and 130 of AGM A3G were found to be adaptive because they confer resistance to Vif-mediated antagonism [98]. In order to further characterize the age and distribution of the genetic conflict between *A3G* and *vif*, we asked whether adaptive signatures in *A3G* were common to a wide range of OWM hosts. Full-length *A3G* was amplified from single representatives of OWM species, including members of the divergent *Colobinae* subfamily. In addition, the dataset was supplemented by previously published sequences from NCBI GenBank [90,133] (Figure 18).

We found that mutations at codons 128 and 130 of *A3G* have emerged independently in several primate lineages (Figure 18). In particular, a D130A mutation was detected in four members of the *Cercopithecus* genus: the De Brazza's monkey (*C. neglectus*), Wolf's guenon (*C. wolfi*), the lesser white nosed monkey (*C. petaurista*), and the mustached guenon (*C. cephus*) (Figure 18). We found D130A homozygosity in three of the four species, while the lesser white nosed monkey is heterozygous. The presence of the D130A mutation in four separate guenon species suggests that it has approached fixation since emerging in the ancestor of the *Cercopithecus* genus 5-6 MYA [17]. In addition to D130A, one allele of A3G from the mustached guenon contains a K128E mutation, a Vif-blocking SNP previously observed in a proportion of grivet monkeys (AGM) (Figure 12) [98].



Species	126	127	128	129	130	131	132	n	source
AGM - Grivet	F	W	**K / E**	P	D	Y	Q	10	[19]
AGM - Sabaeus	F	W	K	P	**D / H**	Y	Q	10	[19]
Patas Monkey	F	W	K	P	D	Y	Q	1	#
Mustached	F	W	**K / E**	P	**A**	Y	Q	1	#
Less White Nosed	F	W	K	P	**D / A**	Y	Q	1	#
Wolf's Guenon	F	W	K	P	**A**	Y	Q	1	#
De Brazza's	F	W	K	P	**A**	Y	Q	1	#
Allen's Swamp	F	W	K	P	D	Y	Q	1	#
Rhesus Macaque	F	W	K	P	**D / N**	Y	Q	36	[28]
Crab Macaque	F	W	K	P	D	Y	Q	1	[29]
Crested Macaque	F	W	K	P	D	Y	Q	1	[29]
Red Mangabey	F	W	K	P	D	Y	Q	1	#
Sooty Mangabey	F	W	K	P	D	Y	Q	1	#
Olive Baboon	F	W	K	P	D	Y	Q	1	#
Francois' Leaf	F	W	K	P	D	Y	Q	1	#
Proboscis Monkey	F	W	K	P	D	Y	Q	1	#
Mantled Colobus	F	W	K	P	D	Y	Q	1	#
Chimpanzee	F	W	**D**	P	D	Y	Q	1	[29]
Bonobo	F	W	**D**	P	D	Y	Q	1	[29]
Human	F	W	**D**	P	D	Y	Q	1	[29]
Gorilla	F	W	**D**	**Q**	D	Y	Q	1	[29]
Orangutan	F	W	**D**	P	D	Y	Q	1	[29]

Figure 18. Diversifying selection of the Vif binding site of *APOBEC3G* in the *Cercopithecinae* subfamily of Old World Monkeys

A partial primate species phylogeny is depicted as a cladogram and approximate divergence times are placed at relevant ancestral nodes [17]. Select residues comprising the putative sites of A3G targeted by Vif proteins are shown. Primates of the order *Catarrhini*, consisting of hominoids (blue) and Old World Monkeys, are included. Old World Monkeys are subdivided into the *Cercopithecinae* (green) and *Colobinae* (red) subfamilies. Sources of sequences previously reported elsewhere are indicated by reference number. # = new sequences reported in this study, n = number of individuals analyzed per species.

Previous work has demonstrated that *A3G* is undergoing adaptive evolution in primates, as measured by the relative rates of non-synonymous variation (dN) and synonymous variation (dS) [90,126]. However, given the additional interspecies and intraspecies *A3G* sequences reported in this study, we reexamined the gene's evolutionary history within the OWM clade to determine if the pattern of non-synonymous mutation is suggestive of selection, and moreover, of exposure to a common selective pressure [105,134] (Figure 19 and Table 2). In agreement with previous efforts, our results indicate that the *A3G* locus as a whole is evolving according to diversifying (positive) selection in OWM (Table 2). A model that allows sites to evolve under selection (M8) provides a significantly better fit to the molecular data than does a model of neutral evolution (M7) (Table 2). Using the mixed-effected model of evolution (MEME) [104] to identify specific residues that are subject to diversifying selection, we find strong signals of selection originating from 13 codons spread throughout the length of the gene, including codons 128 and 130 of the Vif binding site ($p = 0.0496$ and $p = 0.0493$, respectively) (Table 3). The random effects likelihood (REL) analysis also identified codons 128 and 130 as being under positive selection, with both sites displaying $dN/dS > 1$ with high posterior probabilities (> 0.99) (Table 3). Overall, our simulations support that the constellation of mutations identified at codons 128 and 130 of *A3G* result from natural selection, suggesting that a genetic conflict between *A3G* and *vif* is proceeding in multiple species of OWM.

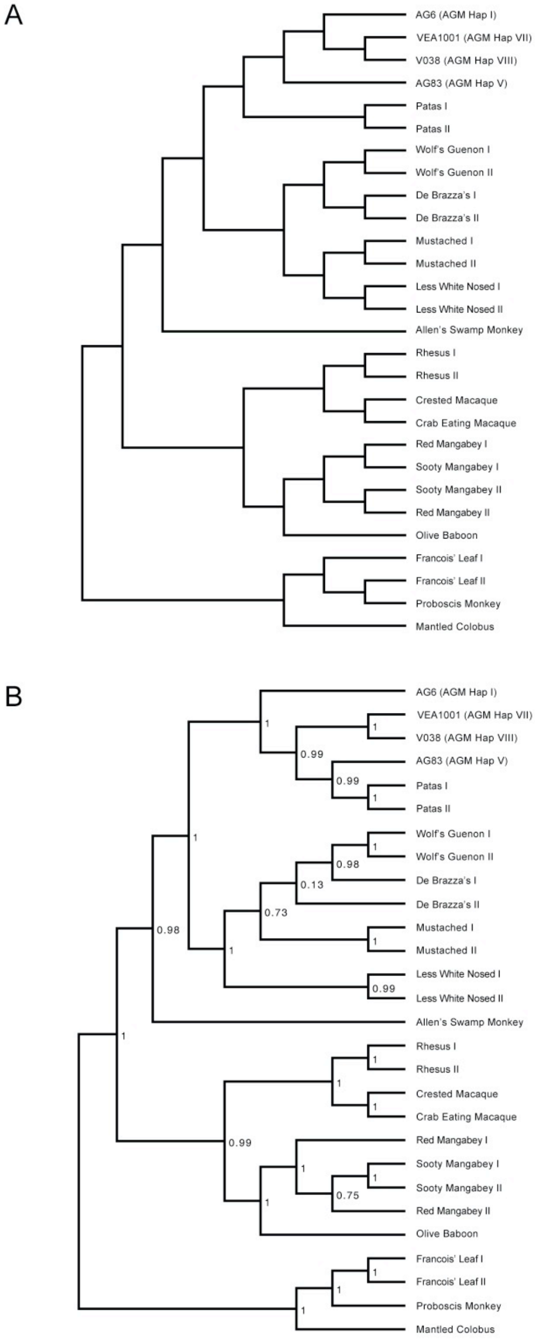


Figure 19. Phylogeny of OWM *A3G*
(A) The phylogeny utilized for evolutionary analyses, modeled after the accepted OWM species phylogeny [17]. (B) A bootstrapped maximum likelihood phylogeny of OWM *A3G* produced using the web-based version of PhyML (phylogeny.fr) and depicted as a cladogram. Note the somewhat different branching topology and paraphyly of some intraspecies *A3G* sequences.

Data Set	Selection Models	df	2(Δln λ)	*p*-value
OWM + Hominoids n = 33	M7 vs M8	2	170.5	$p < 0.001$
	M0 vs M1	63	48.1	$p = 0.664$
OWM only n = 28	M7 vs M8	2	153.4	$p < 0.001$
	M0 vs M1	53	37.8	$p = 0.943$
OWM only Partition aa 78 - 190	M7 vs M8	2	28.7	$p < 0.001$
	M0 vs M1	53	26.2	p = 0.999

Table 2. Comparison of results from selection models implemented in Codeml
A comparison of M7 and M8 models, assuming the F3x4 codon frequency model and an initial ω value of 0.4, was used to identify individual residues of A3G undergoing positive selection. M7 and M8 allow the dN/dS values for each site to vary according to a beta distribution, but M8 allows for dN/dS values greater than 1 (selection). A comparison of M0 (one ratio) and M1 (free ratio) models was used to distinguish between pervasive selection throughout the tree and episodic selection localized to individual branches. The difference in likelihood scores generated by each model was calculated by likelihood ratio tests, using 2(n-1)-1 degrees of freedom. Data sets: OWM + Hominoids, OWM alone, and OWM alone partitioned between recombination breakpoints.

Codon (as in AGM A3G)	*p*-value		Posterior Probability that dN/dS > 1			
	MEME		REL (Bayes Factor)		BEB	
	Full-length	Partitioned	Full-length	Partitioned	Full-length	Partition aa 77 - 143
11	0.0069	0.0069	0.999 (94059)	1.00 (5x10^6)	1.000	--
13	0.0106	0.0164	0.999 (5530)	0.999 (44422)	1.000	--
14	0.0600	0.0570	0.999 (27413)	0.999 (10920)	0.999	--
15	0.0753	0.0729	0.999 (1440)	0.999 (3541)	1.000	--
18	0.0564	0.0517	0.999 (651)	0.999 (2440)	0.999	--
29	0.0090	0.0114	0.999 (69014)	0.999 (332631)	1.000	--
46	0.0200	0.0371	0.992 (36)	0.937 (9)	0.701	--
55	0.0415	0.0411	0.999 (22638)	0.999 (14535)	0.999	--
59	0.0872	0.0913	0.999 (757)	0.998 (316)	0.936	--
75	0.0448	0.0372	0.999 (6515)	0.999 (23057)	1.000	--
128	0.0496	0.0836	0.993 (47)	0.955 (13)	0.856	0.884
130	0.0493	0.0126	0.999 (12736)	0.999 (2395)	0.996	0.998
143	0.0964	0.1032	0.999 (5138)	0.996 (154)	0.999	1.000
168	0.0847	0.0821	0.999 (2290)	0.998 (567)	0.993	--
170	0.0431	0.0300	0.999 (647)	0.999 (1389)	0.990	--
172	0.0167	0.0117	0.994 (47)	0.969 (19)	0.877	--
204	0.0762	0.0673	0.994 (48)	0.990 (63)	0.748	--
227	0.0325	0.0332	0.999 (41)	0.999 (6434)	1.000	--
330	0.0300	0.0281	0.999 (36740)	0.999 (4846)	1.000	--
355	0.0031	0.0182	0.999 (22911)	0.999 (5966)	0.998	--

Table 3. Sites undergoing diversifying selection in OWM *A3G*
The web-based version of HyPhy (datamonkey.org) was used to perform the MEME and REL analyses. The Nsites feature in Codeml was used to perform the BEB analysis. Shown are sites identified by MEME to meet or approach the threshold for diversifying selection ($p < 0.05$). REL and BEB scores for these same sites are reported as posterior probabilities and as Bayes factors (in parentheses).

Ancient, high frequency SNPs in OWM A3G allow escape from Vif-mediated degradation

To test whether the D130A and K128E mutations in A3G are adaptive by affecting susceptibility to Vif-mediated antagonism, we measured the sensitivity of A3G variants to a panel of Vif proteins, including autologous isolates (derived from viruses naturally circulating in a given host species) and heterologous isolates (those derived from other host species). Since viral adaptation can mask the adaptive phenotype of genetic mutations of the host, the use of a broad spectrum of Vif proteins, each with a different adaptive history, allows us to assess how mutations in A3G affect sensitivity to antagonism.

The antiviral activity of A3G from De Brazza's monkey, Wolf's guenon, and mustached guenon was measured by co-expressing each variant with Vif-deficient HIV-1. All A3G variants inhibit infectivity of the virus by more than 100-fold relative to virus produced in the absence of A3G, demonstrating that the antiviral activity of A3G orthologs has been conserved despite the observed variation (Figure 20a).

Using recombinant HIV-1 virus engineered to express SIV *vif* genes, we found that De Brazza's A3G is resistant to antagonism by Vif from SIVagm.Ver, a heterologous isolate that infects vervet monkeys (AGM) (Figure 20a). However, reversion of D130A (A130D, restoration of ancestral ^{128}KPD^{130}) renders De Brazza's A3G fully sensitive, demonstrating that a single alanine residue at position 130 allows escape from antagonism. Conversely, Vif from the virus that naturally circulates among De Brazza's monkeys in the wild, SIVdeb [135], readily antagonizes De Brazza's A3G despite the D130A adaptation. A3G from two other members of the *Cercopithecus* genus

(Mustached guenon (allele I) and Wolf's guenon) exhibits the same sensitivity as De Brazza's A3G (Figure 20a). These data suggest that the emergence of D130A in De Brazza's monkeys (or the *Cercopithecus* ancestor) drove SIVdeb Vif to adapt to this highly prevalent A3G variant, a phenomenon we call 'counter-evolution'. Furthermore, De Brazza's A3G is also sensitive to SIVagm.Sab Vif, which antagonizes A3G carrying D130A or A130D. This activity is likely the result of prior adaptation to the D130H polymorphism in A3G from sabaeus monkeys (AGM), as previously described (Figure 12) [98]. Western blot analysis confirms that antagonism of A3G variants by Vif results in depletion of intracellular A3G protein. De Brazza's A3G expression is reduced substantially in the presence of SIVdeb Vif, relative to expression in the absence of Vif, whereas SIVagm.Ver Vif has no impact (Figure 20b). Upon reversion of the D130A mutation, however, SIVagm.Ver Vif is capable of depleting A3G levels (Figure 20b). These results suggest that the derived D130A mutation was selected 5-6 MYA to evade ancestral SIV Vif proteins.

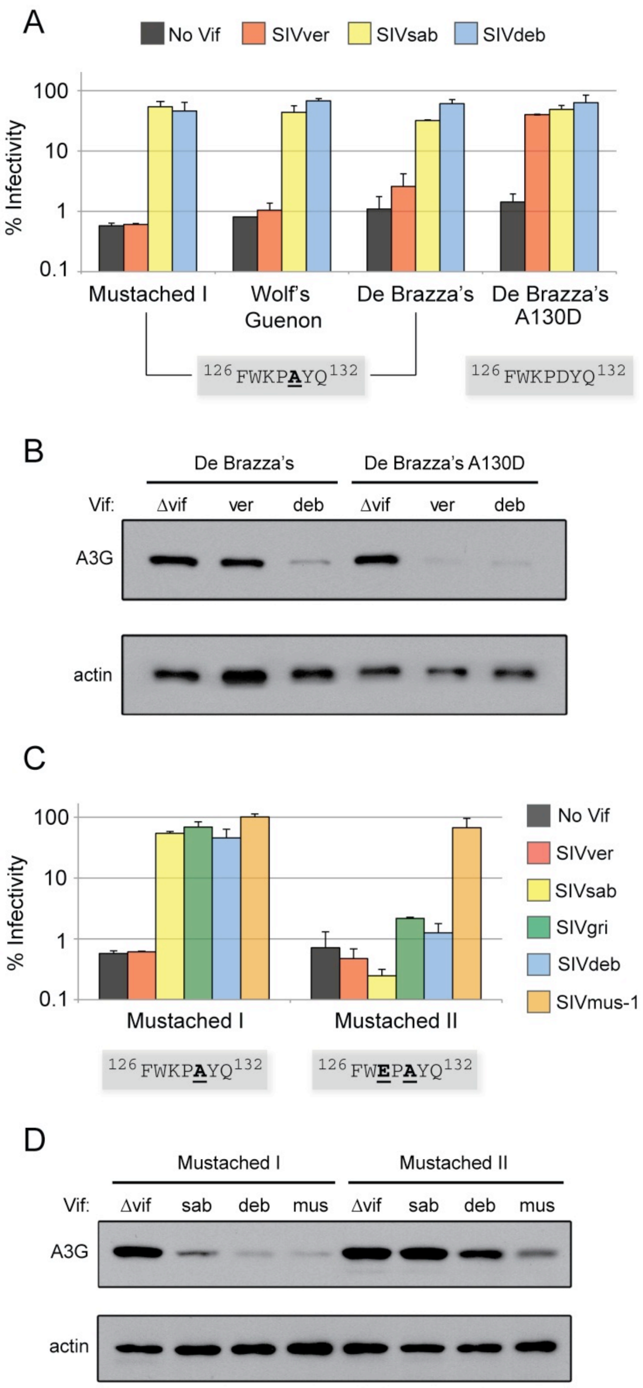


Figure 20. Adaptive evolution of *A3G* at codons 128 and 130 allows escape from Vif-mediated antagonism Single-round infectivity assays were performed with HIV-1ΔVif and HIV-1 expressing SIV Vif proteins produced in the presence of A3G variants from monkeys of the *Cercopithecus* genus in (A) and (C). Infectivity of viruses is reported as a percentage, relative to infectivity in the absence of A3G (100%). Error bars indicate standard deviation from the mean of two independent transfection experiments (six infection replicates in total). The Vif binding site of each A3G variant is depicted in grey boxes. Anti-HA western blot analysis was used to measure A3G expression in virus producer cells in (B) and (D). Anti-ß-actin served as protein loading controls.

Recent virus-driven evolution at residue 128 of A3G

One variant of A3G specific to the mustached guenon, the ^{128}EPA^{130} variant encoding K128E in addition to D130A (Mustached II), is resistant to four heterologous Vif proteins: SIVagm.Ver Vif, SIVagm.Sab Vif, SIVagm.Gri Vif, and SIVdeb Vif isolates (Figure 20c). This demonstrates that K128E, like D130A, prevents Vif-mediated antagonism. However, this variant of A3G is sensitive to Vif from SIVmus-1, indicating that at least one of the three lentivirus strains currently circulating in mustached guenons [43,136] has counter-evolved while adapting to this species (Figure 20c). Concordantly, only SIVmus-1 Vif depletes expression levels of the ^{128}EPA^{130} variant (Mustached II), while the ^{128}KPA^{130} variant (Mustached I) common to other *Cercopithecus* monkeys is degraded by several Vif proteins (Figure 20d). These data demonstrate that K128E and D130A were selected at different times during primate evolution to prevent Vif-mediated antagonism of A3G, with the latter occurring 5-6 MYA in the common ancestor of the *Cercopithecus* genus and the former appearing recently in a single species (the mustached guenon). Recurrent virus-driven evolution of *A3G* over time suggests that natural host species are engaged in a prolonged, antagonistic relationship with lentiviruses.

Broad specificity of some Vif proteins may facilitate cross-species transmission

While most examples of variation at the Vif binding site of *A3G* were identified in natural hosts of modern SIV strains, we also found unique variation among rhesus macaques (*Macaca mulatta*). Captive macaques have experienced outbreaks of simian AIDS stemming from accidental cross-species transmissions of SIVsm (giving rise to SIVmac) in the 1970s [22,137], but they are not thought to harbor a lentivirus in wild Asian habitats. Using a previously published dataset from 36 Indian-origin rhesus macaques, a D130N polymorphism in *A3G* was identified in 59/74 (80%) of chromosomes examined [133] (Figure 18). Like the D130H and D130A mutations observed in sabaeus monkeys and members of the *Cercopithecus* genus, respectively, rhesus A3G encoding D130N (variant ^{128}KPN^{130}) resists antagonism by SIVagm.Ver Vif (Figure 20a). Since SIVsm was able to cross-transmit into both humans and macaques, with both species exhibiting specific variation at the Vif-binding site of A3G (Figure 18), we tested the activity of SIVsm Vif. Similarly to the macaque-adapted strain SIVmac, Vif from SIVsm is capable of antagonizing rhesus A3G despite the D130N mutation (Figure 21a). Western blot analysis demonstrates that both variants of rhesus A3G are depleted by SIVsm Vif, but not by HIV-1 Vif, while SIVagm.Ver Vif is only capable of degrading the variant encoding the ancestral ^{128}KPD^{130} motif (Rhesus I) (Figure 21c). Furthermore, human A3G is susceptible to antagonism by HIV-2 Vif as well as Vif from SIVsm, in agreement with a prior report [138], but not SIVagm.Ver Vif (Figure 21a and 21b). These data demonstrate that SIVsm Vif, in exhibiting broad cross reactivity for the A3G substrate, was ‘pre-optimized’ to target both rhesus A3G and human A3G prior to cross-species transmission. Moreover, this capacity for widespread antagonism has been maintained by SIVmac and HIV-2 following emergence in rhesus macaques and humans, respectively.

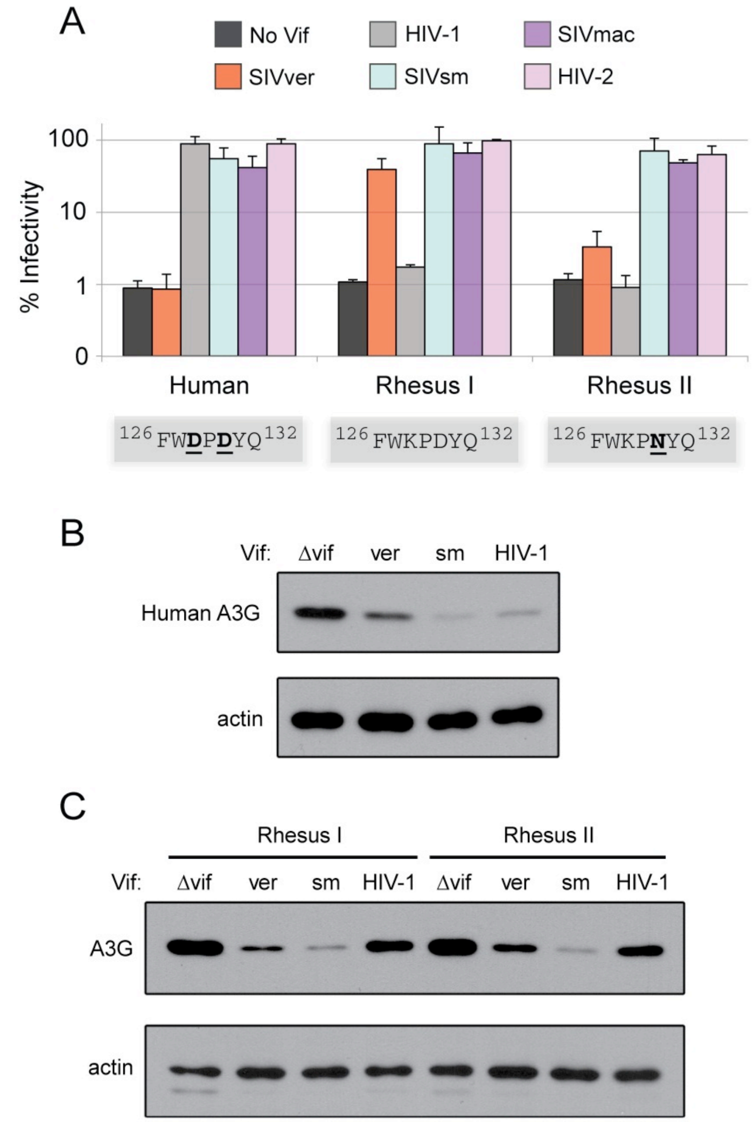


Figure 21. Broad and potent activity of Vif from SIVsm and its descendants allow antagonism of rhesus and human A3G variants

Single-round infectivity assays were performed with HIV-1ΔVif and HIV-1 expressing SIV Vif proteins produced in the presence of A3G variants from human and rhesus macaques (A). Infectivity of viruses is reported as a percentage, relative to infectivity in the absence of A3G (100%). Error bars indicate standard deviation from the mean of two independent transfection experiments (six infection replicates in total). The Vif binding site of each A3G variant is depicted in grey boxes. Anti-HA western blot analysis was used to measure human A3G expression in (B) and rhesus A3G expression in (C). Anti-ß-actin served as protein loading controls.

Our characterization of SIV Vif proteins suggests that some have evolved to tolerate variation at the Vif-binding site of A3G. To determine *vif* counter-evolution produces antagonists that continue to rely on residues 128 and 130 or whether it shifts the stage of the genetic conflict to distinct surfaces on A3G, we tested the activity of Vif proteins against seven A3G variants representing each variation of the Vif-binding site. While the range of A3G variants targeted by each Vif varies, no Vif was capable of antagonizing the full spectrum (Table 4). SIVmus-1 Vif fails to inhibit human A3G and only minimally inhibits the two AGM A3G variants. Furthermore, Vif from SIVagm.Sab recognizes A3G from AGM, rhesus, and De Brazza's, but cannot tolerate ^{128}EPA^{130} present in mustached guenon A3G. The broadest acting Vif species, encoded by SIVsm and its descendants (SIVmac and HIV-2), exhibit specificity for nearly all variants of A3G reported here. However, all three are defective at targeting the ^{128}KPH^{130} variant found in sabaeus monkeys (Table 4). These data indicate that, despite differences in substrate specificity, Vif isolates from viruses infecting *Cercopithecinae* monkeys share a dependency on residues 128 and 130 for antagonism of A3G. Therefore, Vif is most likely the selective agent responsible for the recurrent selection of 'escape' mutations at these positions (Figure 18).

	SIVagm Ver	SIVagm Sab	SIVmus	SIVsm	SIVmac	HIV-2	SIVolc
Rhesus I ^{128}KPD^{130}	39 +/- 16	51 +/- 6	60 +/- 5	89 +/- 59	63 +/- 24	98 +/- 5	< 1
Rhesus II ^{128}KPN^{130}	3 +/- 2	66	59	70 +/- 34	49 +/- 4	63 +/- 19	< 1
De Brazza's ^{128}KPA^{130}	2 +/- 1	32 +/- 1	79 +/- 39	50	90	89 +/- 29	1
Mustached II ^{128}EPA^{130}	< 1	< 1	68 +/- 28	105	52	70	< 1
AGM Sabaeus ^{128}KPH^{130}	< 1	64 +/- 22	8	< 1	17	3 +/- 1	< 1
AGM Grivet ^{128}EPD^{130}	< 1	< 1	5	32	55	83 +/- 26	< 1
Human ^{128}DPD^{130}	< 1	< 1	2	55 +/- 23	42 +/- 18	89 +/- 14	< 1
Colobus ^{128}KPD^{130}	1	3 +/- 1	14 +/- 2	2	1	3 +/- 2	75 +/- 30

Table 4. Sensitivity of A3G variants to a spectrum of diverse SIV Vif proteins
Infectivity of viruses produced in the presence of various A3G variants is reported as a percentage, relative to infectivity in the absence of A3G (100%). Standard deviations from the mean of two to three independent transfection experiments (six to nine infection replicates in total) are reported below. Rows are labeled as OWM A3G variants. Columns are labeled as SIV Vif isolates.

A multi-residue insertion that blocks Vif emerged approximately 12 MYA in A3G of the Colobinae ancestor

In studying the species-specificity of SIV Vif proteins, we found that A3G from the mantled colobus monkey (*Colobus guereza*) is widely resistant to most SIV Vif proteins (Table 1), despite carrying the ancestral ^{128}KPD^{130} at the Vif binding site (Table 4). This observation suggests that residues in A3G lying outside of the canonical Vif binding motif can govern susceptibility to antagonism. The mantled colobus species (hereafter referred to as colobus) belongs to the *Colobinae* subfamily of OWM, a group of primates that diverged from *Cercopithecinae* about 18 MYA [17]. It is naturally associated with a specific SIV strain termed SIVcol, and this is the case for other closely related species (SIVwrc in western red colobus and SIVolc in olive colobus) [139-143]. In testing the sensitivity of colobus A3G to Vif from viruses naturally associated with *Colobinae* hosts, we found that SIVolc Vif was unique in its ability to target it for destruction (Figure 22a and 22b). In fact, SIVolc Vif antagonizes solely colobus A3G and not A3G from any other primate species tested (Table 4). Conversely, SIVagm.Sab Vif exhibits the opposite specificity, readily counteracting AGM A3G and several OWM A3G orthologs but not colobus A3G (Table 4).

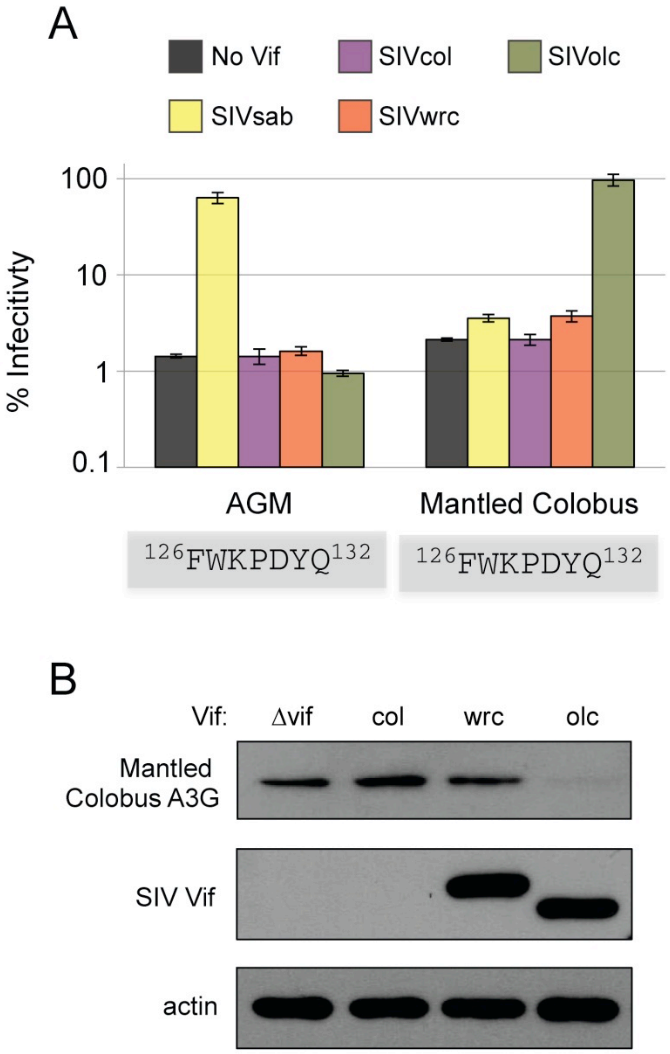


Figure 22. SIVolc Vif antagonizes colobus A3G
(A) Single-round infectivity assays were performed with HIV-1ΔVif and HIV-1 expressing SIV Vif proteins produced in the presence of AGM haplotype I A3G or Colobus A3G. Error bars indicate standard deviation from the mean of three infection replicates. SIVolc Vif, but not Vif from other characterized isolates of SIV that infect *Colobinae* hosts, overcomes restriction by A3G of the mantled colobus (*Colobus guereza*). (B) Anti-HA and anti-FLAG western blot analysis were used to measure expression of A3G and Vif in virus producing cells, respectively. Note that expression of Vif from SIVcol, the species-specific lentivirus of the mantled colobus monkey, was not detected. SIVwrc Vif was expressed, but demonstrated no activity against colobus A3G.

To learn how colobus A3G remains resistant to nearly all Vif proteins except SIVolc Vif, we constructed chimeric A3G proteins containing portions of the N-terminus of AGM A3G and the C-terminus of colobus A3G. These chimeras were co-expressed with virus encoding SIVagm.Sab Vif or SIVolc Vif to test for sensitivity to antagonism. The critical constructs are shown in Figure 22. Chimera C and chimera D differ by only seven amino acids, yet the former is sensitive to SIVagm.Sab Vif while the latter is resistant (Figure 23a). Interestingly, a multi-residue insertion unique to members of the *Colobinae* subfamily is contained within this sequence (Figure 23b). Upon removal of the insertion (^{66}SCK^{68}) from wild-type colobus A3G, a full gain in sensitivity to SIVagm.Sab Vif is achieved (Figure 23a, compare Colobus del 64-66 to Colubus A3G). Therefore, a three amino acid insertion that emerged in the N-terminus of A3G in the *Colobinae* ancestor prevents antagonism by Vif. The Vif-blocking activity of ^{66}SCK^{68} is context dependent in that it only blocks other Vif proteins when in combination with residues 66-199 of colobus A3G (data not shown). Therefore, for it to be selected to evade antagonism by Vif, this 3 amino acid insertion in A3G could only have evolved once during primate evolution—in the *Colobinae* ancestor.

In comparison to SIVagm.Sab Vif, SIVolc Vif displayed different specificities for the same chimeric A3G proteins, demonstrating that it has diverged to target distinct surfaces of the A3G substrate (Figure 23a, compare Chi B to Chi A). Mutagenesis of residues within 110-165 that are divergent between AGM A3G and colobus A3G reveal that E133, N137, K141, and A145 (depicted as 'ENKA') are major recognition determinants of SIVolc Vif (Figure 23c and 23d). Single mutations of E133 or N137 alone completely prevent antagonism by SIVolc Vif, while mutation of K141 and A145 in combination also blocks antagonism (Figure 23d). Interestingly, residues 133, 137, and 145 are divergent between members of the *Colobinae* subfamily, suggesting that this motif may be diversifying in response to Vif from SIV infecting these primates (Figure 23c).

In order to determine if SIVolc Vif antagonizes A3G independently of the "canonical" Vif interaction motif involving residues 128 and 130, we tested its ability to antagonize A3G encoding against naturally occurring mutations at these sites. Indeed, we found that A3G constructs that are sensitive to SIVolc Vif remain so after the introduction of the K128E and D130A mutations (the ^{128}EPA^{130} motif found in mustached A3G) (Figure 23e). Conversely, these mutations completely abrogated antagonism by SIVagm.Sab Vif (Figure 23e). Therefore, SIVolc Vif has diverged to utilize unique surfaces of A3G while adapting to its natural host, targeting residues that are divorced from those targeted by all other Vif proteins studied to date.

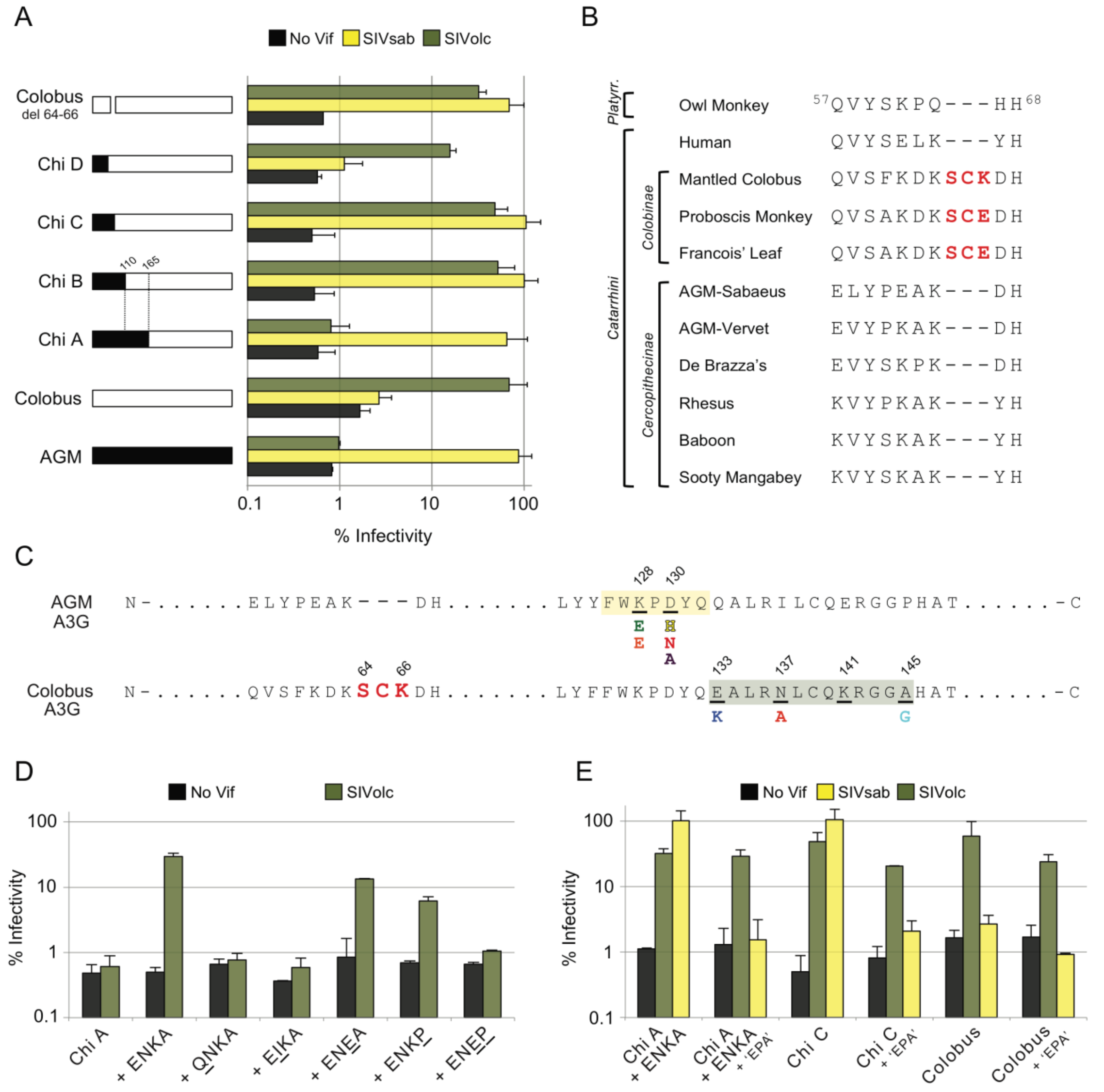


Figure 23. An insertion that blocks Vif emerged 12 MYA in *A3G* of the *Colobinae* ancestor. (A) Single-round infectivity assays were performed with HIV-1ΔVif and HIV-1 expressing SIV Vif proteins produced in the presence of colobus/AGM chimeric A3G proteins. Infectivity of viruses is reported as a percentage, relative to infectivity in the absence of A3G (100%). Error bars indicate standard deviation from the mean of two independent transfection experiments (six infection replicates in total). (B) Partial protein alignment of A3G orthologs reveals a three-residue insertion (red) unique to members of the *Colobinae* subfamily of OWM. Members of primate parvorders *Catarrhini* (OWM and Hominoids) and *Platyrrhini* (New World Monkeys) are included. OWM are further divided into subfamilies *Cercopithecinae* and *Colobinae*. (C) Residues of A3G responsible for differential sensitivity to SIVsab and SIVolc Vif. The canonical Vif binding site is boxed in yellow. Divergent character states independently selected at residues 128 and 130 (underlined) of *Cercopithecinae A3G* are displayed below. The three-residue insertion in the N-terminus of colobus A3G is bolded in red. Residues of A3G required for antagonism by SIVolc Vif, as found in this study, are underlined and boxed in green. Divergent character states at residues 133, 137 and 145 identified in *Colobinae A3G* are displayed below. (D) and (E) Single-

round infectivity assays were performed with HIV-1ΔVif and HIV-1 expressing SIVolc Vif produced in the presence of mutated colobus/AGM chimeras.

Discussion

We have used the approach of studying virus-driven host evolution to discern the minimum age for the association between OWM and SIV at 5-6 MYA, and possibly 12 MYA. Moreover, using a broad range of primates and SIV isolates, we have identified both recurrent and novel interactions between Vif with A3G. Finally, our data suggests that the broad specificity exhibited by some Vif proteins may facilitate cross-species virus transmission events.

Adaptive evolution at the Vif binding site of A3G in Cercopithecinae monkeys is a molecular beacon for past lentivirus infections

In this study, we report that the region of A3G targeted by Vif is independently diversifying in several primate lineages. A previous report of positive selection in primate *A3G* concluded that Vif (and by extension, lentiviruses) did not play a major role in the gene's evolution [90]. However, using a data set enriched with a number of polymorphic natural hosts of SIV, we find charge-altering mutations at residues 128 and 130 of the Vif binding site. Using maximum likelihood methods, we show that these sites are evolving under diversifying selection. Furthermore, *in vitro* infections reveal that single amino acid changes affect sensitivity to Vif-mediated degradation. While we cannot exclude the possibility that this region of *A3G* is subject to selective pressures other than lentiviral Vif, recurrent selection at precisely the same sites targeted by most SIV Vif proteins supports that the selective agent responsible is a Vif-encoding element. These data suggest that Vif drives the emergence of 'escape' mutations in *A3G* that allow evasion of Vif-mediated degradation, which in turn promotes *vif* counter-evolution and the perpetuation of a genetic conflict between host and virus.

What remains unresolved is whether or not the ancient pathogens inferred by this study are the direct ancestors of contemporary SIV strains. We raise two possibilities: 1) extant lentiviruses are themselves ancient, having coexisted continuously with their specific hosts for millions of years, or 2) extant lentivirus infections are young, such that adaptive evolution at the A3G-Vif interface was driven by lentiviruses that no longer exist (paleoviruses). The second scenario posits that modern SIV strains may not necessarily bear semblance to the lentiviruses that drove selection in *A3G*, a distinct possibility given the prevalence of cross species transmission [45], dual lentivirus infections and circulating recombinant forms [37,136,144], and virus lineage extinction [145]. That is, the evolutionary histories of natural host species may be punctuated by periodic lentivirus infections rather than by a single, enduring lentiviral threat.

Along with our previous discovery of adaptive evolution in *A3G* of AGM [98], the emergence of 'escape' mutations in members of the *Cercopithecus* genus yield insight into the age and pathogenic potential of SIV infections in natural host species. The D130A mutation common to the four members of the genus indicates that *Cercopithecus* ancestors were exposed to a form of SIV prior to speciation, one that impacted host fitness and selected for adaptive mutations in innate immunity. Furthermore, the subsequent emergence of E128K within one of those species, the mustached guenon, suggests the selective pressure applied by SIV is not only ancient but also *ongoing*. This particular mutation is unlikely to be found in other members of *Cercopithecus*, since SIVmus-1 is unique in its ability to degrade it. Vif from SIVdeb, a strain infecting a closely related host, the De Brazza's monkey, does not tolerate variation at this site (Figure 20c). Therefore, E128K likely represents a more recent adaptive change than the D130A mutation common to the genus. The sequential emergence of two mutations that each allowed escape from Vif proteins suggests that the mustached guenon lineage has been subjected to continuous (scenario 1, above) or periodic (scenario 2) selective pressure by SIV since diverging from other members of the genus 5-6 MYA. This is considerably older than previous phylogenetic analyses have indicated, and presents an alternative route to dating viral infections that does not suffer from the limitations of virus sequence-based methods [61,62].

Another example of variation in the Vif-binding motif of A3G was found in the rhesus macaque, a species of Asian descent that is not known to carry a circulating lentivirus. Because the SIVmac epidemic is too recent to have imprinted on rhesus immunity, the novel D130N polymorphism specific to this population (Figure 17) may have appeared in response to an SIV Vif of unknown origin, be it extinct or currently uncharacterized. The genetic heterogeneity of rhesus macaques is also evident in the *TRIM5* gene, which is highly polymorphic and gives rise to seven distinct variants with anti-lentivirus activity [70]. Moreover, one allele encodes a TRIM5-CypA fusion protein that restricts HIV-2 and SIVagm, suggesting that it was selected for 5-6 MYA by a virus with similar characteristics [146-149]. Together with the data presented here on *A3G* polymorphism, the rhesus genome abounds with genetic clues alluding to a former lentiviral presence.

Broad specificity of SIV Vif originating from Cercopithecinae hosts with polymorphic A3G

Our results indicate that, upon adaptation to one or more variants of A3G in a polymorphic host species, SIV evolves Vif proteins with broader specificity. That is, by counter-evolving to antagonize resistant variants of A3G from their respective host species, SIV Vif gains the ability to target A3G variants present in other species. This is evidenced by SIVagm.Sab Vif which, having adapted to target ^{128}KPD^{130} and ^{128}KPH^{130} variants of A3G in sabaeus monkeys, exhibits the capacity to antagonize the ^{128}KPA^{130} variant in *Cercopithecus* monkeys and the ^{128}KPN^{130} variant in Rhesus macaques (Figure 20a, Table 4). This is in stark contrast to the related SIVagm.Ver Vif, which solely antagonizes A3G bearing the ^{128}KPD^{130} motif ancestral to OWM. Furthermore, SIVmus-1 Vif has adapted to persist among mustached guenons, which have presented adaptive variation at both residues 128 and 130 of A3G. In doing so, SIVmus-1 Vif can cross-react with a broad array of A3G orthologs that present different combinations of characters at these two sites (Table 4). Likewise, in the cases of SIVmac and HIV-2, the broad activity of their Vif proteins may have been pre-determined by events that played out in sooty mangabey populations (Table 4). Our data suggest that SIVsm was ‘pre-optimized’ to target both rhesus A3G and human A3G prior to cross-species transmission, and that this activity has been maintained by SIVmac and HIV-2 following emergence in rhesus macaques and humans, respectively. Thus, we believe that the cross-reactivity or ‘promiscuity’ of Vif proteins can shine light on the unique adaptive history of lentivirus strains, by providing clues about the *A3G* diversity of its host species.

A novel insertion in Colobus A3G has altered Vif targeting preferences

While the canonical Vif binding site of A3G was conserved in members of the divergent *Colobinae* subfamily of OWM, we found that a multi-residue insertion (^{66}SCK/E^{68}) in the N-terminus renders A3G resistant to nearly all Vif proteins. As of yet there is no crystal structure of the A3G-Vif interaction, but our experimental results demonstrate how the insertion disrupts the ability for Vif to counteract A3G. Upon removal of these three residues from colobus A3G, this variant becomes sensitive to antagonism by SIVagm.Sab Vif, revealing that all the determinants necessary for binding and degradation are intact elsewhere in the protein. Thus, we hypothesize that the ^{66}SCK^{68} insertion serves to conceal the typical Vif binding site by altering protein conformation and masking distal epitopes necessary for Vif-mediated antagonism. Furthermore, our studies using chimeric A3G proteins demonstrate that the insertion functions as such only in the context of colobus A3G. Thus, the multi-residue insertion may represent an alternative strategy to evade antagonism by Vif proteins, one that could only have evolved within the *Colobinae* lineage. Our functional analysis suggests that this adaptive feat drove the evolution of SIV Vif proteins with different targeting preferences. That is, in adapting to the presence of the ^{66}SCK^{68} insertion in colobus A3G, SIVolc Vif has evolved to recognize a patch of residues offset from the binding site preferred by other Vif proteins. Therefore, the stage of the genetic conflict between A3G and Vif has shifted at least once during primate evolution. The consequences of this switch in substrate recognition may already be evolutionarily apparent as residues of A3G important for SIVolc Vif-mediated antagonism are divergent in *Colobinae* species used in this study (Figure 23c).

In summary, the data reported herein allow us to infer the presence of lentiviruses that applied pathogenic selective pressure at different points in primate evolutionary history. In a marked display of convergent evolution, two residues of A3G that coordinate an interaction with SIV Vif are diversifying in multiple primate species. Moreover, a divergent strategy of Vif-evasion has emerged in a separate branch of the primate phylogeny, giving rise to a second genetic conflict and altering the interface of the A3G-Vif interaction. The pattern of adaptive mutation suggests that SIV has been infecting OWM on a million-year timescale.

Chapter V: Perspectives

We conclude that the functional polymorphism and divergence identified in *A3G* is a demonstration of natural selection in species challenged with pathogenic lentiviruses. However, a review of other examples of innate immune variation in primates and its impact on lentivirus infection and disease is necessary to frame our results in a greater context.

Paleovirology: viral artifacts and virus-driven evolution of host immunity

Technological advances in genome sequencing have unearthed endogenous viral elements (EVE), the remnants of ancient viruses that successfully invaded animal genomes. In fact, a range of recent endogenous discoveries has revealed that several virus families with modern relevance to human health have far more distant origins than previously appreciated [4,5]. Support for a long-lived association between primates and lentiviruses was provided by an intact endogenous 'fossil' resembling a lentivirus uncovered in the genomes of Malagasy lemurs. Called prosimian SIV (pSIV), this germ line integration event establishes that lentiviruses were circulating among at least some primates 4 million years ago (MYA) [58,59]. However, pSIV is a transitional viral form, sharing some aspects of genomic architecture with nonprimate lentiviruses (such as *dUTPase*) while lacking genes typically found in primate lentiviruses (such as *vpr*) [58].

Regardless of its true classification, no trace of a lentivirus-like entity has been identified in the genomes of simian primates, the African hosts of modern primate lentiviruses.

In the absence of direct evidence of past infections, viruses can be detected indirectly by examining how primate antiviral immunity has evolved. Discoveries of protective phenotypes that appear to confer resistance to virus infection or to infection-associated pathogenesis suggest that viruses can drive the evolution of the hosts they infect. In a limited but growing number of cases, the genetic underpinnings of such adaptations have been traced to immunity genes in primate genomes. A signature of natural selection, manifesting as point mutations or indels within protein-coding regions of the locus, is common to these adaptive genotypes. The best estimate for selective agents responsible for these selective imprints are pathogenic viruses, those causing death or other significant costs to fitness on the population level. Spontaneous, formerly neutral mutations in antiviral genes will be ‘selected’ at the onset of an emerging epidemic and will spread through the host population at a rate proportional to the severity of the infection. The allelic frequency of genetic variants increases rapidly to the point of absolute or near fixation in the case of a selective sweep (positive selection), or alternatively, mutations may persist for long periods of time at intermediate frequencies (balancing selection).

For example, relative to humans, plasmacytoid dendritic cells (pDCs) of sooty mangabeys express dramatically less type 1 interferon (IFN) in response to viral nucleic acid sensing by Toll-like receptor 7 (TLR7) and TLR9 [150]. Whereas HIV-1 provokes chronic IFN production and widespread immune activation in humans, the response to SIV in sooty mangabeys and AGM is transient and subsides after the acute phase of infection [150,151]. This distinction in signaling outcomes is associated with species–specific amino acid changes in the transcription factor IRF-7, which shuttles into the nucleus to induce interferon-stimulated genes. Therefore, a genetic adaptation in an innate immune signaling molecule of a natural host was likely selected after an extensive association with lentiviruses.

Analysis of standing variation in immunity genes from a broad panel of extant primates, representing tens of millions of years of evolution, provide a temporal framework for when such adaptations emerged in the past. The rate at which non-synonymous mutations accumulate between species (known as divergence) is a classic measure of adaptive evolution, also known as positive selection [152]. The discovery of an extremely diverse set of innate immune genes with lentivirus blocking potential, known as host restriction factors, has accelerated our understanding of how pathogenic viruses impact host evolution [66]. Many restriction factors are evolving in accordance with selection, and in fact, this is one criterion by which they are defined. Furthermore, because many of these factors are directly antagonized by the lentiviruses they have evolved to restrict, instances of past infections by a specific virus family can be inferred when the precise interface between a restriction factor and the viral antagonist is known [153]. In this way, the adaptive evolution of restriction factors in primates can be read like a genetic ledger, documenting when a given host population was exposed to a lentivirus threat.

Our examination of naturally occurring variation in *A3G* from OWM species revealed that lentivirus-driven evolution could be traced through time to specific primate lineages [98] (Chapters III & IV). Because *A3G* is evolving in tandem with the lentiviral accessory gene *vif*, host evolution can be linked to past lentivirus infections in deep and shallow timescales. Sequencing of *A3G* orthologs from a broad panel of OWM, including AGM and several other species comprising the SIV reservoir, revealed recurrent selection at residues involved in the interaction with lentiviral Vif (Figure 18). Single mutations that allow evasion of Vif-mediated degradation independently emerged on several occasions, the oldest of which being 5-6 MYA (Figure 20 and 21). An even older Vif-blocking adaptation in *A3G*, a multi-residue insertion event, arose approximately 12 MYA in the *Colobinae* subfamily of OWM that also harbors lentiviruses (Figure 23). These findings suggest that ancient lentiviruses left footprints of selection stamped in the *A3G* locus.

Similar inferences can be drawn from the evolutionary history of newly discovered restriction factor SAMHD1, which imposes a block to viral reverse transcription in myeloid-derived blood cells. Sequencing the *SAMHD1* locus from representatives of major primate clades revealed that it is undergoing positive selection [154]. The majority of signal originates from the SAM domain, at positions that also affect sensitivity to the lentiviral-encoded antagonists Vpr and Vpx. Interestingly, this selection is most pronounced in members of the *Cercopithecinae* lineage, the OWM subfamily that includes AGM, sooty mangabeys, and a large component of natural host species. While many sites of SAMHD1 may be rapidly evolving, the diversification of residues that are binding determinants for Vpr/Vpx suggests that a lentiviral selective pressure is responsible. These data collectively illustrate that lentiviruses have influenced the evolution of primate hosts for millions of years, most notably those that are associated with lentiviruses today.

‘Natural’ SIV infections are inherently antagonistic and require genetic vigilance

While most examples of lentivirus-driven selection examines interspecies divergence, presumably ‘fixed’ mutations that distinguish the antiviral gene of one species from another, selection can also manifest as variation *within* a species (known as polymorphism). When applied to the evolution of antiviral immunity, polymorphism signals relatively recent instances of virus-driven evolution.

Analysis of gene diversity within AGM, a prototypic and highly abundant natural host species, have turned up troves of informative polymorphism. AGM are comprised of four major subspecies that occupy distinct geographies in sub-Saharan Africa. Known commonly as vervet, tantalus, grivet, and sabaeus monkeys, each population harbors a distinct SIV strain (collectively referred to as SIVagm). In two of those subspecies, single amino acid changes in A3G emerged independently to allow evasion of the lentiviral antagonist, Vif. Specifically, these adaptive mutations prevent degradation by Vif from heterologous SIVagm isolates (those found in the other subspecies), suggesting that they were selected as a result of Vif-mediated selective pressure [98] (Figure 12). Additionally, unique polymorphisms at the same sites of *A3G* were also observed in the mustached guenon (*Cercopithecus cephus*), a species that plays host to three distinct strains of SIV in the wild (Figure 17).

Earlier findings of polymorphism in the chemokine receptor *CCR5* of AGM provide further evidence of lentivirus-driven selection in this natural host species. CCR5 is utilized by the HIV/SIV as a co-receptor to enter host cells. Non-synonymous mutations in *CCR5* of AGM subspecies cluster in regions targeted by the viral Envelope glycoprotein, and said variation inhibits SIVagm infection *in vitro* [155]. Together, signatures of recent selection in *A3G* and *CCR5* demonstrate that both antiviral and proviral host factors are subject to ongoing evolutionary change during the course of host-virus coevolution.

Whereas variation between species allows us to infer the presence of ancient, pathogenic lentiviruses among primate ancestors, selection resolved to the population level hints that contemporary lentiviruses may also impact the fitness of natural hosts. Indeed, cases of AIDS, while rare, have been documented in SIV-infected AGM and sooty mangabeys in captivity [127,128]. Furthermore, cross-species transmission of apparently nonpathogenic SIV strains, both in the laboratory and in the wild, can induce disease in the non-natural host [120,156]. Even in the absence of full-blown AIDS, SIV infections may have subtler, unrecognized consequences. For example, infection may affect reproductive rate or social behaviors such as mate choice. Thus, primate lentiviruses may be inherently poised for pathogenesis, demanding a ‘prepared’ and finely tuned immune reaction from the host.

The prospect of pathogenic lentiviruses among wild monkey populations may have important implications for our understanding of HIV-1 pathogenesis in humans. SIVcpz, the ancestor of HIV-1, is a recombinant lentivirus with a mosaic genome, incorporating elements from viruses that infected ancestors of the red capped mangabey (SIVrcm) and guenons (SIVgsn/SIVmon/SIVmus) [157]. A nine-year field study showed that SIVcpz-infected wild chimpanzees suffer from decreased life span and an AIDS-like illness [130,158]. Therefore, the pathogenic nature of SIVcpz in chimps and, subsequently, HIV-1 in humans, may be an extension of the antagonistic interplay carried out between lentiviruses and *Cercopithecinae* primates. Long-term field studies examining lentivirus transmission and disease prevalence within natural host populations are needed to exclude this possibility.

The ongoing diversification of antiviral genes, as indicated by the emergence of adaptive polymorphism in innate immune factors, may represent a strategy to keep potentially life-threatening lentiviruses at bay. It has been suggested that genetic variability of a host population may impede virus spread and survival, since adaptation to multiple host genotypes may restrict virus evolution [159]. In fact, immune polymorphism may not only benefit populations but also individuals, in the form of heterozygosity. A demonstration of this phenomenon was made possible by examining SIV evolution following experimental infection of monkeys with differing *A3G* genotypes [115]. In a monkey heterozygous for *A3G*, possessing one ancestral A3G and one derived, Vif-resistant variant of A3G, SIVagm *vif* failed to evolve the capacity to antagonize both variants (Figures 14) [98]. In contrast, *vif* readily counteracted A3G in individuals homozygous for either allele, even when antagonism of the Vif-resistant variant necessitated virus evolution (counter-evolution) (Figure 14) [98]. The restraint applied by host heterozygosity is apparent from the fact that adaptive mutations in *vif* that enabled antagonism of derived A3G conferred a loss of activity against the ancestral form (Figures 16 and 17). These data support that heterozygosity can, at least temporarily, provide an adaptive advantage to the host and shift the balance of the genetic conflict.

In addition to describing the immune adaptations of natural host species that have resulted from, and possibly enabled, long-term coexistence with lentiviruses, some observations of primate immunity suggest that lentiviruses may have been more widespread in the past than they are today.

Ebbs and flows of the selective tide: lentivirus extinction and re-emergence

Species-specific SIV strains have been defined in dozens of African non-human primates, yet the relationship between lentiviruses and some species is more ambiguous. Surveys of SIV in the wild have identified ‘spillover’ hosts that are not infected with unique versions of SIV but rather harbor transient infections normally found in other species. Examples include the yellow baboon, the chacma baboon, and the white crowned mangabey, three species that, at low prevalence, carry an SIVagm subtype usually confined to AGM with which they share habitat [40,160,161]. While clearly exposed to lentiviruses today, it is uncertain whether such species were hosts to species-specific strains in the past.

Similarly, the first report of lentivirus infection in patas monkeys, a species that is closely related to and sympatric with sabaeus monkeys (AGM) in Senegal, showed that approximately 5% of animals were cross infected with the SIVagm.Sab subtype [162]. A later study showed that patas monkeys share the immunophenotypic features of canonical natural host species, namely the downregulation of the CD4 receptor from helper T cells, suggesting that a hypothetical patas-specific SIV once existed [145]. This observation raises the possibility that adaptation of host immunity may in fact mediate resistance to infection and drive the extinction of lentivirus strains. A similar conclusion may be drawn for the white-crowned mangabey, which exhibits relatively low frequency of CD4+ T cells, a property shared by other mangabey species that are known natural hosts [163].

The possibility that genetic adaptations may provide resistance to lentivirus infections in the wild, and in turn drive their extinction, provides encouragement to gene therapeutic approaches being developed to control the spread of the HIV-1 pandemic.

While it may be reasonable to suspect that most (if not all) primates in Africa have been infected by lentiviruses at one time or another, primate habitats span the globe from the New World to the Far East. Current surveillance efforts limit the range of SIV infections to the African continent, but a trail of genetic evidence allows one to follow the spread of lentiviruses out of Africa and into far flung primate habitats.

In parallel with the emergence of HIV in humans, Asian macaque species have been burdened by AIDS epidemics of their own, also the consequence of cross-species transmission from sooty mangabeys in captivity [137]. As the primary non-human models of AIDS, the genetic heterogeneity of Asian macaque species has confounded efforts to identify what constitutes an effective immune response against modern pandemics [71,164-167]. However, this variability can be exploited to better understand the age and distribution of primate lentiviruses. For example, innovation in the innate immune repertoire of the rhesus macaque (*Macaca mulatta*) may have been inspired by past and present encounters with lentiviruses. Rhesus macaques encode the well-characterized host restriction factor TRIM5α as well as a novel TRIM-CypA fusion protein, with the latter emerging in ancestral macaques 5-6 MYA [146,148,149]. While TRIM5α potently blocks HIV-1 replication, TRIM-CypA exhibits specificity towards HIV-2 and SIVagm, suggesting that it was selected millions of years ago in response to a virus with similar characteristics [147,168]. In a remarkable demonstration of convergent evolution, a TRIM-CypA fusion has evolved independently in owl monkeys (*Aotus*), a genus endemic to South America, establishing that species of the New World may have been prey to lentiviruses as well [169]. Therefore, evolution and expansion of the *TRIM* arsenal in macaques and other atypical primates implies that SIV may lurk in unexpected places.

The possibility of a natural rhesus-specific SIV strain is reinforced by trademarks of selection imprinted in other host factors. A high-frequency SNP in *A3G* of Indian-origin rhesus macaques, like those identified in other members of the *Cercopithecinae* subfamily of OWM, encodes a unique charge-altering amino acid change that affects susceptibility to the viral antagonist Vif (Figure 21). This polymorphism may represent a vestigial footprint of an extinct lentivirus (paleovirus) that once circulated among rhesus macaques. Otherwise, it is a nascent adaptive mutation signaling the presence of an as-yet-unidentified SIV in Asia. Sequence analysis of single representatives of additional macaque species did not reveal the same or any other SNP at this position of *A3G*, but further examination of *A3G* diversity in Chinese-origin rhesus and other macaque species throughout Asia is warranted.

Perhaps most striking of all, human evolutionary genetics may allude to a lentiviral presence that preceded HIV-1. A unique five amino acid deletion is fixed in the human ortholog of tetherin, a host restriction factor that “tethers” virions to the cell surface to disallow their release and spread. This deletion prevents antagonism of tetherin by the lentiviral Nef protein utilized by many SIV strains, suggesting that it may have been selected for by a pathogenic, Nef-encoding lentivirus approximately 800,000 years ago [80,81,170]. Similarly, some Northern Europeans encode a truncated variant of CCR5 (CCR5Δ32) that mediates resistance to HIV-1 infection [171,172]. While it is difficult to identify the selection pressure responsible, analogous CCR5 truncations can be found in mangabey species. In the sooty mangabey, naturally occurring CCR5Δ2 and CCR5Δ23 variants have driven the evolution of SIV with adapted co-receptor usage [173]. Therefore, adaptive gene reconstructions in innate immune genes of humans may have been triggered by lentivirus infections in our ancestors.

Chapter VI: Future Directions

Identifying the genetic variants responsible for differential outcome of experimental SIV infection in AGM

Our analysis of virus evolution *in vivo* (Figure 14) demonstrates that SIV *vif* can rapidly evolve upon transmission into a foreign host expressing novel variants of A3G. However, based on the viral load and CD4+ T cell count of experimentally infected sabaeus monkeys (reported in [174]), a gain of Vif function was not associated with an increase in peripheral viral load or a decline in CD4+ T cell count. Therefore, A3G is not the sole intracellular barrier restricting viral replication in the setting of cross-species transmission. In an attempt to identify other restriction factors that are variable between the transmitting host (vervets) and the recipient hosts (sabaeus), I found that other restriction factors (*APOBEC3H* and *TRIM5α*) are also polymorphic in the same individuals (data not shown). Subsequent analyses performed by other students in the lab have thus far identified that SAMHD1 and tetherin are also diversifying in a manner that may reflect lentivirus-driven evolution (data not shown). However, when these restriction factor alleles were assembled for each individual, no genetic signatures were associated with resistance to infection. This is likely due to the small sample size of monkeys used in the intial infection study [174]. Therefore, the exteme heterogeneity of these animals complicates our ability to understand the molecular basis for resistance and susceptiblity to natural lentivirus infections. Nonetheless, recent technological advancements in genome sequencing may be utilized to identify all polymorphisms between two populations at once. An RNA-based sequencing platform known as ‘exome’ sequencing will identify all genetic variation present within the coding regions, thus allowing for more efficient and comprehensive SNP discovery. Limiting the sequence reads to transcribed genes allows for reduction of cost and increase in number of samples. Reads could be sorted based on percent nucleotide diversity, allowing one to prioritize gene candidates based on the rate of evolutionary divergence. Furthermore, given a large enough sample size, genomic data could be coupled to experimental infection outcomes or viral loads detected in wild monkeys, allowing one to plumb the genome for genetic determinants of such phenotypes.

Identifying biomarkers of SIV-associated disease in lesser known primate species

Adaptive evolution of the host factor *A3G*, likely the result of lentivirus-mediated selective pressures, identifies that at least some SIV infections of non-human primates may be pathogenic to some extent. Gathering direct evidence of SIV-associated morbidity in wild primate populations is a difficult endeavor; however, phenotypic characterization of a small number of monkeys may supply indirect evidence of pathogenesis. For example, the response of primary cells to *in vitro* virus infections can establish the quality of the innate immune response triggered upon infection. Using the immunophenotypic description of cells derived from AGM and sooty mangabeys as a benchmark, the extent to which the IFN response is resolved following acute infection may help us better classify primate species as natural or non-natural hosts. Therefore, *in* *vitro* infections of cells derived from other putative natural hosts, such as the mustached guenon, will allow us to learn more about the nature of wild SIV infections in this species.

The effect of dual infection and recombination in pathogenic outcomes

Another avenue for future study is to better understand why some host-virus combinations are more likely to cause disease than others. The only SIV strain known to cause overt AIDS in wild populations of its host, SIVcpz, is a recombinant lentivirus that must have emerged in chimpanzees following co-infection by lentiviruses from different sources. Interestingly, we identified two species of OWM that possess Vif-resistant forms of A3G, the sabaeus monkey (AGM) and the mustached guenon, which also are infected by mosaic lentiviruses. It is tempting to speculate that recombination events between two lentivirus strains can give rise to a novel strain and establish host-virus relationships that are ‘out-of-sync.’ That is, a combination of the required host and virus factors needed to resolve immune activation and avoid pathogenesis may be absent upon infection with a novel recombinant lentivirus. In this way, dual lentivirus infections and recombination may potentiate pathogenesis, and cause new epidemics of AIDS. Examination of *A3G* variation is warranted for other species known to harbor multiple lentiviruses, such as the mandrill.

Characterization of naturally-occurring variation in other host restriction factors

Our results, in conjunction with work published elsewhere, suggest that a binary classification of SIV hosts (natural or non-natural) is an oversimplification. There is likely a spectrum of phenotypes that manifest at the molecular, cellular, and organismal level that ultimately relate the extent to which species have been associated with lentiviruses. For example, the patas monkey has no known species-specific SIV infection but exhibits the immunomodulatory adaptations found in AGM and sooty mangabeys, suggestive of a one-time association with lentiviruses in the wild. However, we did not observe any molecular evidence (i.e. adaptive mutations in *A3G*) that suggest lentivirus-driven evolution in this species. Evolutionary analysis of other host restriction factors (TRIM5, SAMHD1, tetherin, and other APOBEC3 family members) is underway and will provide a more comprehensive view on the range of species subjected to lentivirus-driven evolution.

The evolution of Vif function

During the course of my dissertation research, I identified that a vast majority of SIV Vif proteins rely on residues in the ^{126}FWKPDYQ^{132} motif of A3G for antagonism, with one exception. Vif from SIVolc, a lentivirus associated with a member of the *Colobinae* subfamily of OWM, is reliant on a distinct set of residues for antagonism of A3G from its autologous host. While this appears to be a derived trait, it is also possible that SIVolc Vif is utilizing the original, ancestral binding site of A3G, whereas SIV infecting *Cercopithecinae* have themselves diverged to utilize the ‘canonical’ Vif binding site at ^{126}FWKPDYQ^{132}. Indeed, SIVolc and the related SIVcol appear most ancestral to the primate lentiviruses in phylogenetic analysis of the *pol* gene [175]. Examining the activity of Vif encoded by the endogenous pSIV found in Malagasy lemurs, which represents the oldest known *vif* sequence of the primate lentivirus family, will help resolve this question. However, it is possible that pSIV Vif performs ancestral functions that may or may not include A3G antagonism. Experiments using this endogenous isolate may help identify whether Vif fulfills other roles inside the cell, activities that may have been maintained by contemporary lentiviruses.

The site of viral adaptation: transmitter or recipient?

As this dissertation can attest, virus emergence requires overcoming intracellular barriers to virus transmission in an infected individual and subsequent spread between individuals. Due to rate at which host populations diverge genetically over time, especially at antiviral loci that are subject to selection, transmission events occuring between two distinct species (cross-species transmission) are presumably fraught with more barriers that might be encountered between two hosts of the same species. RNA viruses and some retroviruses like HIV are equipped with exceedingly high mutation rates relative to that of the host organism. The mutation rate results from an inherently error-prone polymerase used for virus replication, a defect that is tolerated and, in fact, exploited by the viral quasispecies. A significant but particularly evasive question in the evolutionary biology of viruses is whether pre-existing viral mutations present in the quasispecies facilitate host adaptation. Alternatively, *de novo* mutations arising after entry into the new host organism may provide the adaptive leap that help viruses invade new environments.

Results of my own work (see Chapter III) suggest that viral mutations selected for function in the new host were not already present in the parental virus stock used as the inoculum. That is, basic sequencing and subcloning of the virus inoculum (a ‘quasispecies’ resulting from in vitro virus amplification) did not reveal a pre-existing Y84C polymorphism in *vif* that enabled antagonism of A3G *in vivo*. These results suggest that directed evolution of the virus genome can occur upon introduction into a single individual of a new species. Similar results were obtained when HIV-1 was experimentally introduced into its original chimpanzee host. In this study, the parental HIV-1 virus, which antagonizes tetherin via the *vpu* accessory gene, re-purposed the *nef* gene for this same function after passage within a single chimp host [176]. Two point mutations were found to be responsible for this gain of activity by Nef, which were presumably the result of spontaneous *de novo* mutation and selection once inside the chimanzee cell. However, next-generation sequencing technologies are required to definitively rule out that the adaptive mutations identified in these studies were not already present in the inoculum prior to administration.

On the other hand, there is also substantial evidence that the degree of viral diversity represented in a quasispecies can dictate the survival, persistance, and adaptability of virus popuatlions *in vivo* [177]. Experiments that manipulate the mutation rate of viral polymerase enzymes demonstrate that increased replication fidelity leads to decreased viral fitness, suggesting that the breadth of genotypes represented in the quasispecies positions the virus for future success. However, It is widely believed, and generally supported by experimental evidence, that a severe bottleneck ensues after natural virus transmission *in vivo*. For example, in the case of HIV-1 transmission, viral diversity present during acute infections indicates that normally only a single virion successfully traverses the mucosal membrane. Bottlenecking results in a “founder effect,” leaving just one or few virus particles responsible for establishing infection in the new host. Therefore, it is likely that features of the quasispecies are most important for fitness and survival within a host. Indeed, the appropriate level of viral diversity appears to facilitate virus spread within a host, allowing access to greater number of tissue compartments [178]. Future studies will need to employ experimental infections performed with by high and low does inoculums to control for the effects of quasispecies on host adaptation. Additionally, longitudinal sampling of various tissues during early acute infection and next-gen sequencing of viral RNA will be necessary to understand how single virus particles can acquire the genetic prowess to seed an enduring infection.

Concluding remarks

An appreciation of virus origins may allow us to better predict where they will go in the future, and the trajectory they will follow to get there. Approaches that combine evolutionary genetics with experimental virology will continue to reveal the host-virus interactions co-evolving in lock step, which are likely those that are pivotal to virus infection, transmission, and long-term persistence. The strategies that primates have deployed to cope with lentivirus infections, and an understanding of how such adaptations constrain virus fitness, may be exploited as novel therapeutic strategies. The vast potential for viruses to counter-evolve in face of genetic innovations of their hosts suggests that a combinatorial approach to gene therapy may prove most successful. That is, an enduring blockade of host-virus interactions might only be achieved by introducing genetic changes in multiple proviral and antiviral genes, like those naturally selected in hosts of SIV. Lastly, an appreciation for the size of the lentivirus reservoir, both past and present, and for the pathogenic prowess of the many species-specific virus strains contained therein, will help identify the human populations most at risk of zoonotic infection and virus emergence.

References

1. Holmes E (2011) What Does Virus Evolution Tell Us about Virus Origins? Journal of virology 85: 5247-5251.
2. C. Holmes E (2009) The evolution and emergence of RNA viruses. 254.
3. Salehi-Ashtiani K, Lupták A, Litovchick A, Szostak JW (2006) A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. Science (New York, NY) 313: 1788-1792.
4. Katzourakis A, Gifford RJ, Malik HS (2010) Endogenous Viral Elements in Animal Genomes. PLoS genetics 6: e1001191.
5. Holmes (2011) The evolution of endogenous viral elements. Cell host & microbe 10: 368-377.
6. Dunlap KA, Palmarini M, Varela M, Burghardt RC, Hayashi K, et al. (2006) Endogenous retroviruses regulate periimplantation placental growth and differentiation. Proceedings of the National Academy of Sciences of the United States of America 103: 14390-14395.
7. Dupressoir A, Vernochet C, Bawa O, Harper F, Pierron G, et al. (2009) Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene. Proceedings of the National Academy of Sciences of the United States of America 106: 12127-12132.
8. Locatelli S, Peeters M (2012) Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. AIDS 26: 659-673.
9. Peeters M, Delaporte E (2012) Simian retroviruses in African apes. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 18: 514-520.

10. Mahieux R, Gessain A (2011) HTLV-3/STLV-3 and HTLV-4 viruses: discovery, epidemiology, serology and molecular aspects. Viruses 3: 1074-1090.
11. Lee RV, Prowten AW, Satchidanand SK, Srivastava BI (1985) Non-Hodgkin's lymphoma and HTLV-1 antibodies in a gorilla. N Engl J Med 312: 118-119.
12. Switzer WM, Salemi M, Shanmugam V, Gao F, Cong ME, et al. (2005) Ancient co-speciation of simian foamy viruses and primates. Nature 434: 376-380.
13. Betsem E, Rua R, Tortevoye P, Froment A, Gessain A (2011) Frequent and recent human acquisition of simian foamy viruses through apes' bites in central Africa. PLoS Pathog 7: e1002306.
14. Mouinga-Ondémé A, Caron M, Nkoghé D, Telfer P, Marx P, et al. (2012) Cross-species transmission of simian foamy virus to humans in rural Gabon, Central Africa. J Virol 86: 1255-1260.
15. Boneva RS, Switzer WM, Spira TJ, Bhullar VB, Shanmugam V, et al. (2007) Clinical and virological characterization of persistent human infection with simian foamy viruses. AIDS Res Hum Retroviruses 23: 1330-1337.
16. Wolfe ND, Heneine W, Carr JK, Garcia AD, Shanmugam V, et al. (2005) Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. Proc Natl Acad Sci U S A 102: 7994-7999.
17. Perelman P, Johnson W, Roos C, Seuánez H, Horvath J, et al. (2011) A Molecular Phylogeny of Living Primates. PLoS genetics 7: e1001342.
18. Gallo RC, Salahuddin SZ, Popovic M, Shearer GM, Kaplan M, et al. (1984) Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science 224: 500-503.
19. Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, et al. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220: 868-871.
20. Bouillant AM (1986) Analogy between lymphotropic human retroviruses and large animal retroviruses. Can Vet J 27: 448-454.

21. Clavel F, Guétard D, Brun-Vézinet F, Chamaret S, Rey MA, et al. (1986) Isolation of a new human retrovirus from West African patients with AIDS. Science 233: 343-346.
22. Chakrabarti L, Guyader M, Alizon M, Daniel MD, Desrosiers RC, et al. (1987) Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature 328: 543-547.
23. Guyader M, Emerman M, Sonigo P, Clavel F, Montagnier L, et al. (1987) Genome organization and transactivation of the human immunodeficiency virus type 2. Nature 326: 662-669.
24. Huet T, Cheynier R, Meyerhans A, Roelants G, Wain-Hobson S (1990) Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature 345: 356-359.
25. Hirsch VM, Olmsted RA, Murphey-Corb M, Purcell RH, Johnson PR (1989) An African primate lentivirus (SIVsm) closely related to HIV-2. Nature 339: 389-392.
26. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, et al. (1999) Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature 397: 436-441.
27. Chen Z, Luckay A, Sodora DL, Telfer P, Reed P, et al. (1997) Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. Journal of virology 71: 3953-3960.
28. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, et al. (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science 313: 523-526.
29. Sharp PM, Hahn BH (2011) Origins of HIV and the AIDS Pandemic. Cold Spring Harbor perspectives in medicine 1: a006841.
30. Vallari A, Holzmayer V, Harris B, Yamaguchi J, Ngansop C, et al. (2011) Confirmation of putative HIV-1 group P in Cameroon. J Virol 85: 1403-1407.
31. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, et al. (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455: 661-664.

32. da Silva ZJ, Oliveira I, Andersen A, Dias F, Rodrigues A, et al. (2008) Changes in prevalence and incidence of HIV-1, HIV-2 and dual infections in urban areas of Bissau, Guinea-Bissau: is HIV-2 disappearing? AIDS 22: 1195-1202.

33. Hodges-Mameletzis I, De Bree GJ, Rowland-Jones SL (2011) An underestimated lentivirus model: what can HIV-2 research contribute to the development of an effective HIV-1 vaccine? Expert review of anti-infective therapy 9: 195-206.

34. Chen Z, Luckay A, Sodora, Telfer P, Reed P, et al. (1997) Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. Journal of Virology 71: 3953-3960.

35. Chahroudi A, Bosinger S, Vanderford T, Paiardini M, Silvestri G (2012) Natural SIV Hosts: Showing AIDS the Door. Science (New York, NY) 335: 1188-1193.

36. Peeters M, Courgnard V (Overview of Primate Lentiviruses and Their Evolution in Non-Human Primates in Africa) 2003.

37. Jin MJ, Hui H, Robertson DL, Müller MC, Barré-Sinoussi F, et al. (1994) Mosaic genome structure of simian immunodeficiency virus from west African green monkeys. The EMBO journal 13: 2935-2947.

38. Allan JS, Short M, Taylor ME, Su S, Hirsch VM, et al. (1991) Species-specific diversity among simian immunodeficiency viruses from African green monkeys. Journal of virology 65: 2816-2828.

39. Wertheim JO, Worobey M (2007) A challenge to the ancient origin of SIVagm based on African green monkey mitochondrial genomes. PLoS Pathogens 3: e95.

40. Jin MJ, Rogers J, Phillips-Conroy JE, Allan JS, Desrosiers RC, et al. (1994) Infection of a yellow baboon with simian immunodeficiency virus from African green monkeys: evidence for cross-species transmission in the wild. Journal of virology 68: 8454-8460.

41. Van Heuverswyn F, Li Y, Neel C, Bailes E, Keele BF, et al. (2006) Human immunodeficiency viruses: SIV infection in wild gorillas. Nature 444: 164.

42. Souquière S, Bibollet-Ruche F, Robertson DL, Makuwa M, Apetrei C, et al. (2001) Wild Mandrillus sphinx are carriers of two types of lentivirus. J Virol 75: 7086-7096.
43. Aghokeng AF, Bailes E, Loul S, Courgnaud V, Mpoudi-Ngolle E, et al. (2007) Full-length sequence analysis of SIVmus in wild populations of mustached monkeys (Cercopithecus cephus) from Cameroon provides evidence for two co-circulating SIVmus lineages. Virology 360: 407-418.
44. Aghokeng AF, Ayouba A, Mpoudi-Ngole E, Loul S, Liegeois F, et al. (2010) Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions. Infect Genet Evol 10: 386-396.
45. Charleston MA, Robertson DL (2002) Preferential host switching by primate lentiviruses can account for phylogenetic similarity with the primate phylogeny. Systematic biology 51: 528-535.
46. Sodora DL, Allan JS, Apetrei C, Brenchley J, Douek DC, et al. (2009) Toward an AIDS vaccine: lessons from natural simian immunodeficiency virus infections of African nonhuman primate hosts. Nature Medicine 15: 861-865.
47. Gifford RJ (2011) Viral evolution in deep time: lentiviruses and mammals. Trends in genetics : TIG: 1-12.
48. Merson M, Padian N, Coates TJ, Gupta GR, Bertozzi SM, et al. (2008) Combination HIV prevention. Lancet 372: 1805-1806.
49. Cichutek K, Norley S (1993) Lack of immune suppression in SIV-infected natural hosts. AIDS (London, England) 7 Suppl 1: S25-35.
50. Rey-Cuillé MA, Berthier JL, Bomsel-Demontoy MC, Chaduc Y, Montagnier L, et al. (1998) Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. Journal of virology 72: 3872-3886.
51. Chakrabarti LA, Lewin SR, Zhang L, Gettie A, Luckay A, et al. (2000) Normal T-cell turnover in sooty mangabeys harboring active simian immunodeficiency virus infection. Journal of virology 74: 1209-1223.

52. Hirsch V (2004) What can natural infection of African monkeys with simian immunodeficiency virus tell us about the pathogenesis of AIDS? AIDS reviews 6: 40-53.
53. Diop OM, Gueye A, Dias-Tavares M, Kornfeld C, Faye A, et al. (2000) High levels of viral replication during primary simian immunodeficiency virus SIVagm infection are rapidly and strongly controlled in African green monkeys. Journal of virology 74: 7538-7547.
54. Broussard SR, Staprans SI, White R, Whitehead EM, Feinberg MB, et al. (2001) Simian immunodeficiency virus replicates to high levels in naturally infected African green monkeys without inducing immunologic or neurologic disease. Journal of virology 75: 2262-2275.
55. Zahn R, Rett M, Korioth-Schmitz B, Sun Y, Buzby A, et al. (2008) Simian Immunodeficiency Virus (SIV)-Specific CD8+ T-Cell Responses in Vervet African Green Monkeys Chronically Infected with SIVagm. Journal of virology 82: 11577-11588.
56. Wertheim JO, Worobey M (2009) Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. PLoS computational biology 5: e1000377.
57. Sharp PM, Bailes E, Gao F, Beer BE, Hirsch VM, et al. (2000) Origins and evolution of AIDS viruses: estimating the time-scale. Biochemical Society transactions 28: 275-282.
58. Gifford RJ, Katzourakis A, Tristem M, Pybus OG, Winters M, et al. (2008) A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. Proceedings of the National Academy of Sciences of the United States of America 105: 20362-20367.
59. Gilbert C, Maxfield D, Goodman S, Feschotte C, Malik HS (2009) Parallel Germline Infiltration of a Lentivirus in Two Malagasy Lemurs. PLoS genetics 5: e1000425.
60. Worobey M, Telfer P, Souquière S, Hunter M, Coleman CA, et al. (2010) Island biogeography reveals the deep history of SIV. Science (New York, NY) 329: 1487.

61. Wertheim JO, Kosakovsky Pond SL (2011) Purifying selection can obscure the ancient age of viral lineages. Mol Biol Evol 28: 3355-3365.

62. Holmes EC (2003) Molecular clocks and the puzzle of RNA virus origins. J Virol 77: 3893-3897.

63. Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, et al. (2009) Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS Pathog 5: e1000437.

64. Malim MH, Bieniasz PD (2012) HIV Restriction Factors and Mechanisms of Evasion. Cold Spring Harbor perspectives in medicine 2: a006940.

65. Malim MH, Emerman M (2008) HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell host & microbe 3: 388-398.

66. Duggal NK, Emerman M (2012) Evolutionary conflicts between viruses and restriction factors shape immunity. Nature reviews Immunology.

67. Lemey P, Rambaut A, Pybus OG (2006) HIV evolutionary dynamics within and among hosts. AIDS Rev 8: 125-140.

68. Stremlau M, Perron M, Lee M, Li Y, Song B, et al. (2006) Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. Proceedings of the National Academy of Sciences of the United States of America 103: 5514-5519.

69. Sawyer SL, Wu LI, Emerman M, Malik HS (2005) Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. Proc Natl Acad Sci U S A 102: 2832-2837.

70. Newman RM, Hall L, Connole M, Chen GL, Sato S, et al. (2006) Balancing selection and the evolution of functional polymorphism in Old World monkey TRIM5alpha. Proceedings of the National Academy of Sciences of the United States of America 103: 19134-19139.

71. Kirmaier A, Wu F, Newman RM, Hall LR, Morgan JS, et al. (2010) TRIM5 suppresses cross-species transmission of a primate immunodeficiency virus and selects for emergence of resistant variants in the new species. PLoS biology 8.

72. Neil SJ, Eastman SW, Jouvenet N, Bieniasz PD (2006) HIV-1 Vpu promotes release and prevents endocytosis of nascent retrovirus particles from the plasma membrane. PLoS Pathog 2: e39.
73. Neil SJ, Zang T, Bieniasz P (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature 451: 425-430.
74. Van Damme N, Goff D, Katsura C, Jorgenson RL, Mitchell R, et al. (2008) The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. Cell Host Microbe 3: 245-252.
75. Jouvenet N, Neil SJ, Zhadina M, Zang T, Kratovac Z, et al. (2009) Broad-spectrum inhibition of retroviral and filoviral particle release by tetherin. J Virol 83: 1837-1844.
76. Kaletsky RL, Francica JR, Agrawal-Gamse C, Bates P (2009) Tetherin-mediated restriction of filovirus budding is antagonized by the Ebola glycoprotein. Proc Natl Acad Sci U S A 106: 2886-2891.
77. Mansouri M, Viswanathan K, Douglas JL, Hines J, Gustin J, et al. (2009) Molecular mechanism of BST2/tetherin downregulation by K5/MIR2 of Kaposi's sarcoma-associated herpesvirus. J Virol 83: 9672-9681.
78. Jia B, Serra-Moreno R, Neidermyer W, Rahmberg A, Mackey J, et al. (2009) Species-specific activity of SIV Nef and HIV-1 Vpu in overcoming restriction by tetherin/BST2. PLoS Pathog 5: e1000429.
79. Zhang F, Wilson SJ, Landford WC, Virgen B, Gregory D, et al. (2009) Nef proteins from simian immunodeficiency viruses are tetherin antagonists. Cell Host Microbe 6: 54-67.
80. Sauter D, Schindler M, Specht A, Landford WN, Münch J, et al. (2009) Tetherin-driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV-1 strains. Cell Host Microbe 6: 409-421.
81. Lim E, Malik H, Emerman M (2010) Ancient adaptive evolution of Tetherin shaped Vpu and Nef functions in human immunodeficiency virus and primate lentiviruses. Journal of virology: 1-11.

82. Sheehy AM, Gaddis NC, Choi JD, Malim MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 418: 646-650.

83. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, et al. (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 424: 99-103.

84. Bishop KN, Verma M, Kim EY, Wolinsky SM, Malim MH (2008) APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS Pathogens 4: e1000231.

85. LaRue, Jónsson, Silverstein, Lajoie, Bertrand, et al. (2008) The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. BMC molecular biology 9: 104.

86. Conticello SG, Harris R, Neuberger MS (2003) The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. Current biology : CB 13: 2009-2013.

87. Russell RA, Wiegand HL, Moore MD, Schäfer A, McClure MO, et al. (2005) Foamy virus Bet proteins function as novel inhibitors of the APOBEC3 family of innate antiretroviral defense factors. Journal of virology 79: 8724-8731.

88. Derse, Hill, Princler, Lloyd, Heidecker (2007) Resistance of human T cell leukemia virus type 1 to APOBEC3G restriction is mediated by elements in nucleocapsid. Proceedings of the National Academy of Sciences of the United States of America 104: 2915-2920.

89. Kolokithas, Rosenke, Malik, Hendrick, Swanson, et al. (2010) The glycosylated Gag protein of a murine leukemia virus inhibits the antiretroviral function of APOBEC3. Journal of Virology 84: 10933-10936.

90. Sawyer SL, Emerman M, Malik HS (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS biology 2: E275.

91. Ortiz M, Guex N, Patin E, Martin O, Xenarios I, et al. (2009) Evolutionary Trajectories of Primate Genes Involved in HIV Pathogenesis. Molecular biology and evolution.

92. Bogerd HP, Doehle BP, Wiegand HL, Cullen BR (2004) A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor. Proceedings of the National Academy of Sciences of the United States of America 101: 3770-3774.
93. Huthoff H, Malim MH (2007) Identification of amino acid residues in APOBEC3G required for regulation by human immunodeficiency virus type 1 Vif and Virion encapsidation. Journal of virology 81: 3807-3815.
94. Schröfelbauer B, Senger T, Manning G, Landau NR (2006) Mutational alteration of human immunodeficiency virus type 1 Vif allows for functional interaction with nonhuman primate APOBEC3G. Journal of virology 80: 5984-5991.
95. Xu, Svarovskaia, Barr, Zhang, Khan, et al. (2004) A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance to HIV-1 virion infectivity factor-induced depletion. Proceedings of the National Academy of Sciences of the United States of America 101: 5652-5657.
96. Mariani, Chen, Schröfelbauer, Navarro, König, et al. (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. Cell 114: 21-31.
97. Emerman M, Malik HS (2010) Paleovirology--modern consequences of ancient viruses. PLoS biology 8: e1000301.
98. Compton AA, Hirsch V, Emerman M (2012) The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. Cell host & microbe 11: 91-98.
99. Han G, Worobey M (2012) Endogenous Lentiviral Elements in the Weasel Family (Mustelidae). Molecular Biology and Evolution: 1-4.
100. Regier DA, Desrosiers RC (1990) The complete nucleotide sequence of a pathogenic molecular clone of simian immunodeficiency virus. AIDS Res Hum Retroviruses 6: 1221-1231.
101. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, et al. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36: W465-469.

102. Tosi AJ, Melnick DJ, Disotell TR (2004) Sex chromosome phylogenetics indicate a single transition to terrestriality in the guenons (tribe Cercopithecini). Journal of human evolution 46: 223-237.
103. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD (2006) Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol 23: 1891-1901.
104. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, et al. (2012) Detecting individual sites subject to episodic diversifying selection. PLoS Genet 8: e1002764.
105. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24: 1586-1591.
106. Hahn BH, Shaw GM, De Cock KM, Sharp PM (2000) AIDS as a zoonosis: scientific and public health implications. Science (New York, NY) 287: 607-614.
107. Holmes EC (2008) Evolutionary history and phylogeography of human viruses. Annual review of microbiology 62: 307-328.
108. Koning FA, Newman EN, Kim EY, Kunstman KJ, Wolinsky SM, et al. (2009) Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. Journal of virology 83: 9474-9485.
109. Vetter M, Johnson M, Antons A, Unutmaz D, D'Aquila R (2009) Differences in APOBEC3G expression in CD4+ T helper lymphocyte subtypes modulate HIV-1 infectivity. PLoS Pathogens 5: e1000292.
110. Miyagi E, Brown CR, Opi S, Khan M, Goila-Gaur R, et al. (2010) Stably expressed APOBEC3F has negligible antiviral activity. Journal of virology 84: 11067-11075.
111. Goila-Gaur R, Strebel K (2008) HIV-1 Vif, APOBEC, and intrinsic immunity. Retrovirology 5: 51.
112. Mariani R, Chen D, Schröfelbauer B, Navarro F, König R, et al. (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. Cell 114: 21-31.
113. Xu H, Svarovskaia ES, Barr R, Zhang Y, Khan MA, et al. (2004) A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance

to HIV-1 virion infectivity factor-induced depletion. Proceedings of the National Academy of Sciences of the United States of America 101: 5652-5657.

114. Wolfheim J (1983) Primates of the World: Distribution, Abundance, and Conservation. Seattle, WA: University of Washington Press.

115. Goldstein S, Brown CR, Ourmanov I, Pandrea I, Buckler-White A, et al. (2006) Comparison of simian immunodeficiency virus SIVagmVer replication and CD4+ T-cell dynamics in vervet and sabaeus African green monkeys. Journal of virology 80: 4868-4877.

116. Goldstein S, Ourmanov I, Brown CR, Plishka R, Buckler-White A, et al. (2005) Plateau levels of viremia correlate with the degree of CD4+-T-cell loss in simian immunodeficiency virus SIVagm-infected pigtailed macaques: variable pathogenicity of natural SIVagm isolates. Journal of virology 79: 5153-5162.

117. Müller-Trutwin MC, Corbet S, Tavares MD, Hervé VM, Nerrienet E, et al. (1996) The evolutionary rate of nonpathogenic simian immunodeficiency virus (SIVagm) is in agreement with a rapid and continuous replication in vivo. Virology 223: 89-102.

118. Pandrea I, Apetrei C, Dufour J, Dillon N, Barbercheck J, et al. (2006) Simian immunodeficiency virus SIVagm.sab infection of Caribbean African green monkeys: a new model for the study of SIV pathogenesis in natural hosts. Journal of virology 80: 4858-4867.

119. Fomsgaard A, Hirsch VM, Allan JS, Johnson PR (1991) A highly divergent proviral DNA clone of SIV from a distinct species of African green monkey. Virology 182: 397-402.

120. Hirsch VM, Dapolito G, Johnson PR, Elkins WR, London WT, et al. (1995) Induction of AIDS by simian immunodeficiency virus from an African green monkey: species-specific variation in pathogenicity correlates with the extent of in vivo replication. Journal of virology 69: 955-967.

121. Soares MA, Robertson DL, Hui H, Allan JS, Shaw GM, et al. (1997) A full-length and replication-competent proviral clone of SIVAGM from tantalus monkeys. Virology 228: 394-399.

122. Huthoff H, Malim M (2007) Identification of Amino Acid Residues in APOBEC3G Required for Regulation by Human Immunodeficiency Virus Type 1 Vif and Virion Encapsidation. Journal of virology 81: 3807-3815.

123. An P, Bleiber G, Duggal P, Nelson G, May M, et al. (2004) APOBEC3G genetic variants and their influence on the progression to AIDS. Journal of virology 78: 11070-11076.

124. Do H, Vasilescu A, Diop G, Hirtzig T, Heath SC, et al. (2005) Exhaustive genotyping of the CEM15 (APOBEC3G) gene and absence of association with AIDS progression in a French cohort. The Journal of Infectious Diseases 191: 159-163.

125. Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular biology and evolution 16: 1114-1116.

126. Ortiz M, Bleiber G, Martinez R, Kaessmann H, Telenti A (2006) Patterns of evolution of host proteins involved in retroviral pathogenesis. Retrovirology 3: 11.

127. Traina-Dorge V, Blanchard J, Martin L, Murphey-Corb M (1992) Immunodeficiency and lymphoproliferative disease in an African green monkey dually infected with SIV and STLV-I. AIDS research and human retroviruses 8: 97-100.

128. Ling B, Apetrei C, Pandrea I, Veazey RS, Lackner AA, et al. (2004) Classic AIDS in a sooty mangabey after an 18-year natural infection. Journal of virology 78: 8902-8908.

129. Pandrea I, Silvestri G, Apetrei C (2009) AIDS in african nonhuman primate hosts of SIVs: a new paradigm of SIV infection. Current HIV research 7: 57-72.

130. Keele BF, Jones JH, Terio KA, Estes JD, Rudicell RS, et al. (2009) Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. Nature 460: 515-519.

131. Isbell LA, Cheney DL, Seyfarth RM (1993) ARE IMMIGRANT VERVET MONKEYS, CERCOPITHECUS-AETHIOPS, AT GREATER RISK OF MORTALITY THAN RESIDENTS. Animal Behaviour 45: 729-734.

132. Mangeat B, Turelli P, Liao S, Trono D (2004) A single amino acid determinant governs the species-specific sensitivity of APOBEC3G to Vif action. The Journal of biological chemistry 279: 14481-14483.

133. Weiler A, May GE, Qi Y, Wilson N, Watkins DI (2006) Polymorphisms in eight host genes associated with control of HIV replication do not mediate elite control of viral replication in SIV-infected Indian rhesus macaques. Immunogenetics 58: 1003-1009.

134. Delport W, Poon AF, Frost SD, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26: 2455-2457.

135. Bibollet-Ruche F, Bailes E, Gao F, Pourrut X, Barlow KL, et al. (2004) New simian immunodeficiency virus infecting De Brazza's monkeys (Cercopithecus neglectus): evidence for a cercopithecus monkey virus clade. Journal of virology 78: 7748-7762.

136. Liegeois F, Boue V, Mouacha F, Butel C, Mve-Ondo B, et al. (2012) New STLV-3 strains and a divergent SIVmus strain identified in non-human primate bushmeat in Gabon. Retrovirology 9: 28.

137. Apetrei C, Lerche NW, Pandrea I, Gormus B, Silvestri G, et al. (2006) Kuru experiments triggered the emergence of pathogenic SIVmac. AIDS (London, England) 20: 317-321.

138. Gaddis NC, Sheehy AM, Ahmad KM, Swanson CM, Bishop KN, et al. (2004) Further investigation of simian immunodeficiency virus Vif function in human cells. Journal of virology 78: 12041-12046.

139. Courgnaud V, Pourrut X, Bibollet-Ruche F, Mpoudi-Ngole E, Bourgeois A, et al. (2001) Characterization of a novel simian immunodeficiency virus from guereza colobus monkeys (Colobus guereza) in Cameroon: a new lineage in the nonhuman primate lentivirus family. Journal of virology 75: 857-866.

140. Liégeois F, Lafay B, Formenty P, Locatelli S, Courgnaud V, et al. (2009) Full-length genome characterization of a novel simian immunodeficiency virus lineage (SIVolc) from olive Colobus (Procolobus verus) and new SIVwrcPbb strains from Western Red Colobus (Piliocolobus badius badius) from the Tai Forest in Ivory Coast. Journal of virology 83: 428-439.
141. Courgnaud V, Formenty P, Akoua-Koffi C, Noe R, Boesch C, et al. (2003) Partial molecular characterization of two simian immunodeficiency viruses (SIV) from African colobids: SIVwrc from Western red colobus (Piliocolobus badius) and SIVolc from olive colobus (Procolobus verus). Journal of virology 77: 744-748.
142. Locatelli S, Lafay B, Liegeois F, Ting N, Delaporte E, et al. (2008) Full molecular characterization of a simian immunodeficiency virus, SIVwrcpbt from Temminck's red colobus (Piliocolobus badius temminckii) from Abuko Nature Reserve, The Gambia. Virology 376: 90-100.
143. Locatelli S, Liegeois F, Lafay B, Roeder AD, Bruford MW, et al. (2008) Prevalence and genetic diversity of simian immunodeficiency virus infection in wild-living red colobus monkeys (Piliocolobus badius badius) from the Taï forest, Côte d'Ivoire SIVwrc in wild-living western red colobus monkeys. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 8: 1-14.
144. Takehisa J, Harada Y, Ndembi N, Mboudjeka I, Taniguchi Y, et al. (2001) Natural infection of wild-born mandrills (Mandrillus sphinx) with two different types of simian immunodeficiency virus. AIDS Res Hum Retroviruses 17: 1143-1154.
145. Apetrei C, Gaufin T, Gautam R, Vinton C, Hirsch V, et al. (2010) Pattern of SIVagm infection in patas monkeys suggests that host adaptation to simian immunodeficiency virus infection may result in resistance to infection and virus extinction. J Infect Dis 202 Suppl 3: S371-376.
146. Dietrich EA, Jones-Engel L, Hu SL (2010) Evolution of the Antiretroviral Restriction Factor TRIMCyp in Old World Primates. PloS one 5: e14019.

147. Ylinen LM, Price AJ, Rasaiyaah J, Hué S, Rose NJ, et al. (2010) Conformational adaptation of Asian macaque TRIMCyp directs lineage specific antiviral activity. PLoS Pathogens 6: e1001062.

148. Wilson SJ, Webb BL, Ylinen LM, Verschoor E, Heeney JL, et al. (2008) Independent evolution of an antiviral TRIMCyp in rhesus macaques. Proceedings of the National Academy of Sciences of the United States of America 105: 3557-3562.

149. Newman RM, Hall L, Kirmaier A, Pozzi LA, Pery E, et al. (2008) Evolution of a TRIM5-CypA splice isoform in old world monkeys. PLoS Pathogens 4: e1000003.

150. Mandl JN, Barry AP, Vanderford TH, Kozyr N, Chavan R, et al. (2008) Divergent TLR7 and TLR9 signaling and type I interferon production distinguish pathogenic and nonpathogenic AIDS virus infections. Nature Medicine 14: 1077-1087.

151. Jacquelin B, Mayau V, Targat B, Liovat A, Kunkel D, et al. (2009) Nonpathogenic SIV infection of African green monkeys induces a strong but rapidly controlled type I IFN response. The Journal of clinical investigation 119: 3544-3555.

152. Li W (1997) Molecular evolution. 487.

153. Emerman M, Malik HS (2010) Paleovirology--modern consequences of ancient viruses. PLoS biology 8: e1000301.

154. Lim ES, Fregoso OI, McCoy CO, Matsen FA, Malik HS, et al. (2012) The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx. Cell host & microbe 11: 194-204.

155. Kuhmann SE, Madani N, Diop OM, Platt EJ, Morvan J, et al. (2001) Frequent substitution polymorphisms in African green monkey CCR5 cluster at critical sites for infections by simian immunodeficiency virus SIVagm, implying ancient virus-host coevolution. Journal of virology 75: 8449-8460.

156. Beer BE, Brown CR, Whitted S, Goldstein S, Goeken R, et al. (2005) Immunodeficiency in the absence of high viral load in pig-tailed macaques

infected with Simian immunodeficiency virus SIVsun or SIVlhoest. J Virol 79: 14044-14056.

157. Sharp PM, Shaw GM, Hahn BH (2005) Simian immunodeficiency virus infection of chimpanzees. J Virol 79: 3891-3902.
158. Rudicell RS, Holland Jones J, Wroblewski EE, Learn GH, Li Y, et al. (2010) Impact of simian immunodeficiency virus infection on chimpanzee population dynamics. PLoS Pathog 6: e1001116.
159. Meyerson NR, Sawyer SL (2011) Two-stepping through time: mammals and viruses. Trends in microbiology 19: 286-294.
160. van R, Engelbrecht S, Mwenda J, Laten, Robson, et al. (1998) Simian immunodeficiency viruses (SIVs) from eastern and southern Africa: detection of a SIVagm variant from a chacma baboon. The Journal of general virology 79 (Pt 7): 1809-1814.
161. Tomonaga K, Katahira J, Fukasawa M, Hassan MA, Kawamura M, et al. (1993) Isolation and characterization of simian immunodeficiency virus from African white-crowned mangabey monkeys (Cercocebus torquatus lunulatus). Archives of virology 129: 77-92.
162. Bibollet-Ruche F, Galat-Luong A, Cuny G, Sarni-Manchado P, Galat G, et al. (1996) Simian immunodeficiency virus infection in a patas monkey (Erythrocebus patas): evidence for cross-species transmission from African green monkeys (Cercopithecus aethiops sabaeus) in the wild. The Journal of general virology 77 (Pt 4): 773-781.
163. Vinton C, Klatt N, Harris L, Briant J, Sanders-Beer B, et al. (2011) CD4-Like Immunological Function by CD4- T Cells in Multiple Natural Hosts of Simian Immunodeficiency Virus. Journal of Virology 85: 8702-8708.
164. Cagliani R, Fumagalli M, Biasin M, Piacentini L, Riva S, et al. (2010) Long-term balancing selection maintains trans-specific polymorphisms in the human TRIM5 gene. Human Genetics.
165. Dietrich EA, Jones-Engel L, Hu SL (2010) Evolution of the Antiretroviral Restriction Factor TRIMCyp in Old World Primates. PloS one 5: e14019.

166. Siddiqui R, Krawczak M, Platzer M, Sauermann U, Gay N (2011) Association of TLR7 Variants with AIDS-Like Disease and AIDS Vaccine Efficacy in Rhesus Macaques. PloS one 6: e25474.

167. Nomura T, Matano T (2012) Association of MHC-I genotypes with disease progression in HIV/SIV infections. Front Microbiol 3: 234.

168. Dietrich EA, Brennan G, Ferguson B, Wiseman RW, O'Connor D, et al. (2011) Variable prevalence and functional diversity of the antiretroviral restriction factor TRIMCyp in Macaca fascicularis. J Virol 85: 9956-9963.

169. Sayah DM, Sokolskaja E, Berthoux L, Luban J (2004) Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. Nature 430: 569-573.

170. Sauter D, Vogl M, Kirchhoff F (2011) Ancient origin of a deletion in human BST2/Tetherin that confers protection against viral zoonoses. Hum Mutat 32: 1243-1245.

171. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, et al. (1996) Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. Science 273: 1856-1862.

172. Hummel S, Schmidt D, Kremeyer B, Herrmann B, Oppermann M (2005) Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons. Genes Immun 6: 371-374.

173. Riddick NE, Hermann EA, Loftin LM, Elliott ST, Wey WC, et al. (2010) A novel CCR5 mutation common in sooty mangabeys reveals SIVsmm infection of CCR5-null natural hosts and efficient alternative coreceptor use in vivo. PLoS Pathog 6: e1001064.

174. Goldstein S, Brown CR, Ourmanov I, Pandrea I, Buckler-White A, et al. (2006) Comparison of simian immunodeficiency virus SIVagmVer replication and CD4+ T-cell dynamics in vervet and sabaeus African green monkeys. Journal of virology 80: 4868-4877.

175. Lim ES, Fregoso OI, McCoy CO, Matsen FA, Malik H, et al. (2012) The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx. Cell host & microbe 11: 194-204.

176. Götz N, Sauter D, Usmani SM, Fritz JV, Goffinet C, et al. (2012) Reacquisition of Nef-mediated tetherin antagonism in a single in vivo passage of HIV-1 through its original chimpanzee host. Cell Host Microbe 12: 373-380.

177. Bordería AV, Stapleford KA, Vignuzzi M (2011) RNA virus population diversity: implications for inter-species transmission. Curr Opin Virol 1: 643-648.

178. Gnädig NF, Beaucourt S, Campagnola G, Bordería AV, Sanz-Ramos M, et al. (2012) Coxsackievirus B3 mutator strains are attenuated in vivo. Proc Natl Acad Sci U S A 109: E2294-2303.

Curriculum Vitae

Alex Compton
422 Yale Ave N #407, Seattle, WA 98109
acompton@uw.edu

RESEARCH INTERESTS

Antiviral innate immunity, host-pathogen coevolution, viral zoonotic infections, surveillance of emerging pathogens

EDUCATION

University of Washington; Seattle, WA USA

Ph.D. in Molecular and Cellular Biology **2008 - 2012**

Advisor: Michael Emerman Ph.D, Fred Hutchinson Cancer Research Center
GPA: 3.82
Dissertation: "The evolutionary dynamics of host-pathogen interactions reveal ancient, pathogenic viral infections"
Graduation Date: **November 2012**

Willamette University; Salem, OR USA

B.A. in Biology **2002-2006**

Minors: Chemistry & Spanish
GPA: 3.91, *Magna Cum Laude*
Research Advisor: Barbara Stebbins-Boaz, Department of Biology
Thesis: "Hormonal herbicides used in agriculture disrupt the maturation of amphibian gametes"

La Pontificia Universidad Católica de Valparaiso; Chile

Latin American Studies **2005**

Concentration: Chilean Literature and Poetry

FUNDING & AWARDS

Pasteur Foundation Postdoctoral Fellowship ($70K/yr, 3 yrs)	**2012**
Young Investigator Award, 19th Conference on Retroviruses & Opportunistic Infections	**2012**
National Science Foundation Graduate Research Fellowship ($30K/yr, 3 yrs)	**2010-2012**
National Institutes of Health Training Grant in Viral Pathogenesis ($26K/yr, 1 yr)	**2009-2010**
National Institutes of Health Predoctoral Intramural Research Training Award	**2006-2007**
Phi Beta Kappa Society; Willamette University	**2006**

PUBLICATIONS

Compton A and Emerman M (2012) "Taming the beast: immune adaptation reveals the age, distribution, and pathogenic potential of primate lentiviruses." *In preparation*

Compton A and Emerman M (2012) "Evolution of the APOBEC3G-Vif interaction suggests ancient origins of Simian Immunodeficiency Viruses." *Under review*

Compton A, Hirsch V, and Emerman M (2012) "The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict." ***Cell Host & Microbe*** 11:91-98

Migueles S, Osborne C, **Compton A** et al. (2008) "Lytic granule loading of CD8+ T cells is required for HIV-infected cell elimination associated with immune control." ***Immunity*** 29(6):845-847

LaChapelle A, Ruygrok M, **Compton A** et al. (2007) "The hormonal herbicide 2,4-dichlorophenoxyacetic acid inhibits Xenopus oocyte maturation by targeting translational and post-translational mechanisms." ***Reproductive Toxicology*** 23(7):20-31

CONFERENCE PRESENTATIONS

Compton A and Emerman M (2012) "Antagonism-driven evolution of primate *APOBEC3G* is a footprint of ancient lentivirus infections." *New Perspectives on Immunity to Infection; Heidelberg, Germany.* **Oral presentation**

Compton A and Emerman M (2012) "Detecting signatures of adaptive evolution in *APOBEC3G* to infer ancient, pathogenic forms of SIV in diverse primate taxa." *19th Conference on Retroviruses and Opportunistic Infections (CROI); Seattle, WA.* **Oral themed discussion and poster presentations**

Compton A, Hirsch V, and Emerman M (2011) "Diversification of *APOBEC3G* in African Green Monkeys drives the evolution of SIVagm." *Retroviruses; Cold Spring Harbor Laboratories, NY.* **Oral presentation**

Compton A and Emerman M (2010) "SIVagm Vif exhibits broad specificity that is fine tuned to reflect host-specific adaptation." *West Coast Retrovirus Meeting; Palm Springs, CA.* **Oral presentation**

RESEARCH EXPERIENCE

Post-doctoral Fellow of the **Pasteur Foundation** in the laboratory of **Olivier Schwartz** *Virus & Immunity Unit, Institut Pasteur* Project Title: Dissecting the role of SAMHD1 in lentiviral sensing and innate immune signaling	**2013-**
Doctoral Research in Virology and Molecular Biology in the laboratory of **Michael Emerman** *Human Biology Division, Fred Hutchinson Cancer Research Center* Project Title: Detecting adaptive signatures in primate host restriction factors to reveal the age and pathogenesis of Simian Immunodeficiency Virus	**2008-2012**
Post-baccalaureate Research Fellow in the laboratory of **Anthony Fauci** *Laboratory of Immunoregulation, National Institute of Allergies and Infectious Disease, NIH* Project Title: The functional potential of anti-HIV CD8+ T cell responses of Long-term nonprogressors (LTNP)	**2006-2007**
Science Collaborative Research Program Fellow in the laboratory of **Barbara Stebbins-Boaz** *Department of Biology, Willamette University* Project Title: Characterizing the reproductive consequences of hormonal herbicide exposure in *Xenopus laevis* oocytes	**2004**

TEACHING & MENTORING EXPERIENCE

Lab Mentor, Summer Undergraduate Research Program*, Fred Hutchinson Cancer Center* Supervised and guided a 10-week research project with Julia Behnen, Scripps College Project Title: "Antagonistic viral protein adapts to counteract multiple host factors"	**Summer 2011**
Teaching Assistant, Biochemistry 442: Molecular & Cellular Biology*, University of Washington* Led a discussion section for an advanced biochemistry course for senior undergraduates Devised a weekly quiz and prepared lectures reinforcing key concepts from class	**Fall 2009**
Lab Mentor, BioQuest Academy*, Seattle Biomedical Research Institute* Designed and implemented a research project for secondary school students Introduced students to lab techniques: bacteria culture, RNA isolation, microarray analysis Project title: "The quest for new drug targets against latent tuberculosis"	**Summer 2009**

RELEVANT LABORATORY SKILLS

Virology
- Propagation of Biosafety Level 3 viruses (HIV-1 and SIV), single-cycle and spreading infections in T cells
- Virus culture from clinical specimens (human and primate whole blood, plasma, and PBMC)

Molecular Biology & Biochemistry
- Nucleic acid isolation from clinical specimens (human and primate mononuclear cells)
- qRT-PCR and PCR genotyping, molecular cloning and mutagenesis
- Western Blot analysis and ELiSA

Phylogenetics
- Neighbor-joining (ClustalX), Maximum-Likelihood (PhyML), and Bayesian (BEAST) approaches
- Molecular evolutionary analysis with Data Monkey, HyPhy, and PAML

Flow Cytometry
- Intracellular Cytokine Staining (ICS), cellular proliferation, and targeted cell-killing assays