Ancient loss of APOBEC3H activity in a clade of primates infected with SIV

Erin I. Garcia

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

Doctor of Philosophy

University of Washington
2018

Reading Committee:
Michael Emerman, Chair
Deborah Fuller
Jairam Lingappa

Program Authorized to Offer Degree:
Microbiology

University of Washington

Abstract

Ancient loss of APOBEC3H activity in a clade of primates infected with SIV

Erin I. Garcia

Chair of the Supervisory Committee:
Affiliate Professor Michael Emerman
Microbiology

Restriction factors are an important component of the innate immune response against primate lentiviruses like human immunodeficiency virus (HIV) and simian immunodeficiency viruses (SIV). While they inhibit various steps of the viral replication cycle, lentiviruses have consequently developed strategies to antagonize this restriction. Studying the co-evolution of host restriction factors and their viral antagonists can reveal mechanisms behind viral adaptation to host species. Furthermore, such study can also allow us to appreciate the utility of encoding a diverse repertoire of restriction factors and observe how these loci have evolved over evolutionary time. Genes in the *APOBEC3* family encode cytidine deaminases that can provide a barrier against both retroviruses and retroelements. These enzymes are packaged into budding virions and hypermutate viral genomes upon infection of a new cell. Of all *APOBEC3* genes in humans, *APOBEC3H* (*A3H*) is the most polymorphic. Some haplotypes encode stable and active A3H proteins, while others are unstable and poorly antiviral. Such variation in human A3H affects interactions with the lentiviral antagonist Vif, which counteracts A3H via proteasomal degradation. While *vif* can evolve to antagonize stable and active A3H, not all HIV-1 strains encode Vif proteins that are capable of recognizing and degrading

A3H. In order to broaden our understanding of A3H-Vif interactions, as well as its evolution in Old World monkeys, I characterized A3H variation within four African green monkey (AGM) subspecies, which are natural hosts of SIV. I found that A3H is highly polymorphic in AGMs and has independently lost antiviral activity in multiple Old World monkeys. This loss of function was partially related to protein expression levels but was also influenced by amino acid mutations in the N-terminus. Lower packaging efficiency was additionally observed in AGM A3H after its divergence from the most recent common ancestor with patas monkeys. Moreover, I demonstrate that the evolution of A3H in the primate lineages leading to AGMs was not driven by Vif. This work suggests that activity of A3H is evolutionarily dynamic and may have a negative effect on host fitness, resulting in its recurrent loss in primates.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First, I would like to thank my mentor, Michael Emerman, for his endless optimism and patience throughout my PhD. Your mentorship was a huge factor in my completion of this degree. It was incredible to see how this project transformed through all the Monday morning meetings, lab meetings, and Thursday morning virus meetings over the last few years. I learned how to be more confident in myself (and my data), which will have a huge influence on my career moving forward. I greatly appreciate your ability to target my strengths and help me navigate my future. Thank you for all the lessons, for supporting my goals, and the grueling bike rides.

Thank you to my committee: Jesse Bloom, Deb Fuller, Jason Smith, Jai Lingappa, and Jess Hamerman. Special thanks to Jason for chairing both my exams and Deb and Jai for reading my thesis. Thank you all for the support and advice regarding my science and perseverance towards the finish line.

A huge thank you goes out to my fellow lab members, past and present. The post docs: Ferdi and Nick; our research tech, lab manager, and staff scientist: Louisa, Abby, and Molly; my fellow grad students: Rossana, Mollie, and now Vanessa; and my previous lab mates: Jolien, Oliver, Cristina, Lily (who trained me in the HIV lab), and all the undergrads and rotation students. Thank you for the laughs, advice, 5pm music parties, pranks, soccer matches, and great adventures. I am so thankful I landed in the Emerman lab with such great friends and colleagues. Special thanks to Rossana and Mollie – we were always in this crazy graduate school ride together and you two were the best partners I could have had.

Thank you to my Microbiology cohort – it is such a great program to be in and I'm lucky to have met such fantastic people. In particular, I'd like the recognize the girls in my year: Hannah L, Hannah T, and Ari. I'm glad we became such good friends over the last few years and I can't wait to cheer you on through our careers! Agua Verde will always have a special place in my heart.

There isn't enough space to list everyone but thank you to all my friends outside of lab and Seattle. Knowing that I have a horde of people cheering me on got me through all those times in graduate school where I couldn't see the end of the road.

I'd also like to thank my family. First, my parents, Ann and Andy Garcia. My drive for learning started with you, so thank you for your unconditional support and love throughout all my graduate school whims. I appreciate the emphasis you both put on my education – I never would have gotten this far without you. Also, thank you for the care packages and taking care of my dog. My brother, Craig, for supporting my craziness and making me laugh when I needed it most. Thanks for all the sarcasm and fake fights.

My in-laws: Cecilia Carrion Nieves and Angel de Jesus Crespo. Thank you for welcoming me into your family with open arms and loving me like a true daughter.

Last, but not least, I want to thank my husband, Steven de Jesus Carrion. There are not enough words in the world to describe how much I appreciate you putting up with all the stresses of graduate school a second time around. I'm lucky to have had you by my side through it all. Thank you for being my best friend, cheerleader, and compass during the last four years. You made sure I kept my eye on the prize and that I had confidence in myself.

Dedication

This thesis is dedicated to my husband, Steven

Without your support, this thesis would not exist.

Chapter 1. Introduction

Viruses are ubiquitous in nature, in the air we breathe and our planet’s oceans. Every living thing on Earth, from animals to bacteria, can be infected by a virus. Viruses themselves, though their classification as living or non-living is under debate, can further be infected by virophages. As obligate intracellular parasites, they are dependent on a host for production and can ultimately have negative effects on that host. The first virus discovered, Tobacco mosaic virus, was identified in the late 1800s due to discolored and mottled patterns observed on tobacco leaves [1, 2]. Since then, thousands of viruses have been characterized and our understanding of their biology has led to revelations regarding the development of vaccines and gene therapies. Using viruses, scientists have also made enormous strides to understand cellular processes, genetics, and the evolution of life [3].

To date, more than 220 virus species, including those that are capable of causing disease, are known to infect humans [4]. Humans come into contact with a wide range of viruses on a daily basis, many of which are nonpathogenic to humans. Bacteriophages, for example, are found in human intestines and can infect the bacteria composing a portion of our microbiome, but are incapable of replicating in human cells [5]. Additionally, other viruses with human tropism, like human pegivirus, have no known association with any disease [6]. Nonpathogenic viruses may be advantageous to human health. Bacteriophages can help regulate the human microbiome [7] and human pegivirus has been associated with beneficial effects during infection with other viruses [6].

However, humans are susceptible to numerous pathogenic viruses as well. The first defined pathogenic human virus was yellow fever virus, discovered in 1900 [10]. Since then, hundreds of viruses associated with disease have been described, such as Ebola virus, dengue virus, and polio. Some infections are quickly suppressed by the immune system before disease can manifest, while others go on to cause illness. Though many viruses, such as hepatitis B virus, are thought to share a long evolutionary history with humans [8, 9], several are newly emerging.

Certain viruses, like polio, circulate only in humans. Others are also present in animal reservoirs, like Ebola and dengue virus. When viruses from animal reservoirs cross the species barrier and infect humans, they are called zoonoses or zoonotic viruses. Transmission of zoonotic viruses from animal reservoirs to humans often requires viral adaptation to barriers against infection to support replication and transmission in new hosts. For instance, adaptation of severe active respiratory syndrome coronavirus (SARS-CoV) to the host receptor ACE2 in humans and palm civets allowed for cross-species transmission of the virus from bats to cause the 2002 – 2003 SARS outbreak [11-13]. Additionally, avian influenza virus polymerases require adaptive mutations to overcome defects in viral replication upon infection of human cells [14]. Influenza viruses also encode an interferon antagonist, NS1, that binds to host protein TRIM25 in a species-specific manner, which may play a role in host adaptation [15]. Similarly, Marburg virus protein VP40 and dengue virus protein NS5 both antagonize components of the interferon signaling pathway in humans but cannot counteract alleles found in mice [16, 17], although these have not been investigated in the context of cross-species transmission.

Viruses are best adapted to counteract defenses from their natural hosts, though the culmination of many factors, including viral diversity, can result in cross-species transmission. Cross-species transmissions, while rare, can ultimately lead to novel disease outbreaks in humans that may spread around the globe and place countless lives at risk.

Origins of Human Immunodeficiency Virus

One prominent example of a cross-species transmission with a substantial impact on human health is the human immunodeficiency virus (HIV) pandemic. More than 30 years after the initial discovery of HIV, it still poses a significant public health threat. According to the World Health Organization, 36.7 million people around the world are currently living with the virus, with 1.8 million new infections and 1 million deaths in 2016 alone. While all countries are affected by the pandemic, the greatest disease burden is in sub-Saharan Africa, which accounts for a large portion of cases (Figure 1).

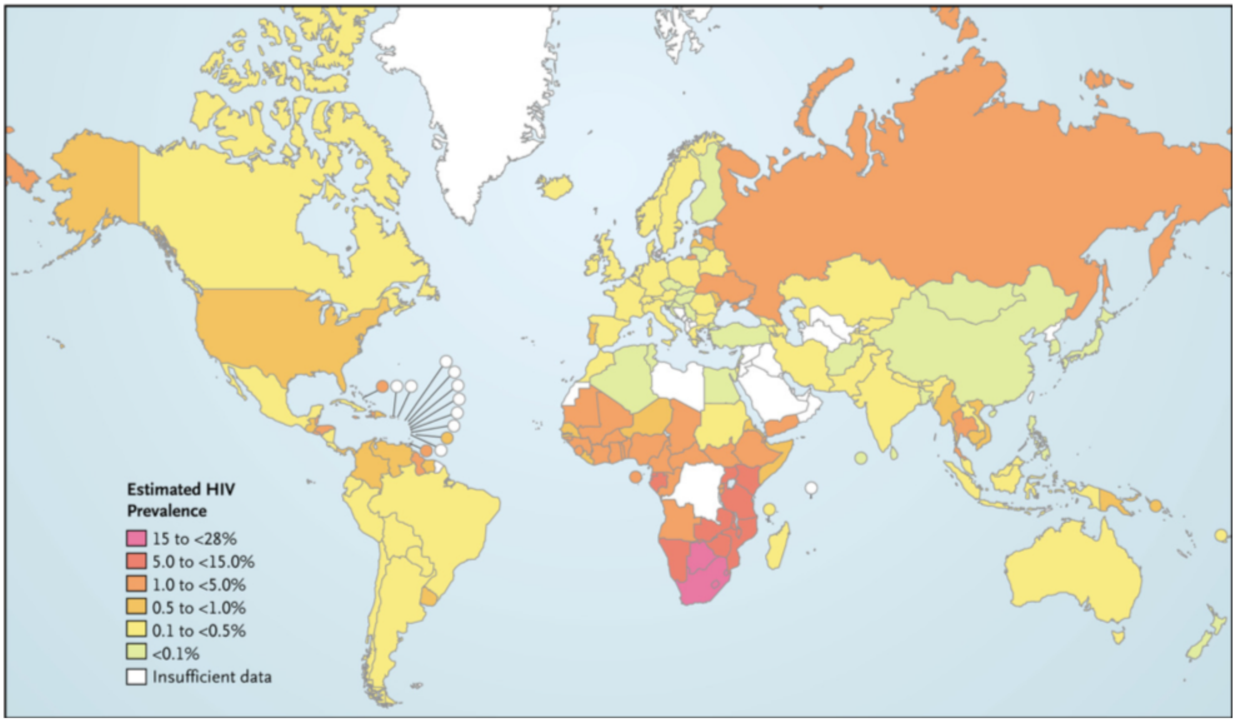

Figure 1. World Map of Prevalence of HIV Infection
Data are from UNAIDS, UNICEF (www.unicef.org), and the World Bank (www.worldbank.org). Reproduced with permission from [18], copyright Massachusetts Medical Society.

HIV belongs to the lentivirus genus of the virus family *Retroviridae*, a group of viruses that reverse transcribe viral RNA into DNA and integrate into host genomes. Lentiviruses infect a broad range of mammals, including sheep, horses, cattle, felines, and primates, to cause chronic and deadly diseases (Figure 2 and Figure 3; [19]). HIV, for instance, preferentially infects adaptive immune cells, particularly CD4+ T-cells [20]. The virus can also exist silently, hidden from host immune responses for an extended period of time after viral integration into the genome of long-lived cells, such as memory T-cells. If left untreated by antiretroviral therapies, chronic HIV infection causes wide-spread immune cell death and progressive failure of immune defenses, called acquired immunodeficiency syndrome (AIDS).

At the beginning of the 1980s, increasing numbers of young homosexual men in the United States were diagnosed with malignancies rarely seen in immunocompetent individuals. Patients who had abused intravenous drugs or received blood transfusions were later also diagnosed with this new syndrome. AIDS appeared to be a long-lasting disease, beginning years after exposure to the agent via blood or during sex. Patients were burdened with an onslaught of opportunistic infections and cancer, tied to a drastic decrease in protective CD4+ T cells. In late 1983, the retrovirus HIV, initially called lymphoadenopathy-associated virus, was identified as the causative agent [21-23]. This led to an explosion of research to characterize the virus and its origins, its interactions with hosts, and to develop treatments. To date, effective prophylactic or therapeutic vaccines have still not been developed, although the implementation of current antiretroviral treatments has vastly improved the lifespan of HIV+ patients [20].

While the first cases of HIV were described in the U.S., it was soon after discovered in African populations [24-26]. The genetic diversity of viruses isolated from Africa has exceeded those from other continents and form phylogenetic outgroups to these strains, suggesting that the origins of the pandemic lay in Central Africa [27-29]. The presence of an AIDS-like disease in rhesus macaques (SIVmac) further led to the discovery that HIV infections had emerged in humans as a consequence of simian immunodeficiency virus (SIV) cross-species transmission from African primates (Figure 2; [30-32]). Two types of HIV circulate in humans, HIV-1 and HIV-2, although HIV-1 has infected the most people worldwide. Collectively, HIV-1 is comprised of four distinct groups: M, N, O, and P, each of which arose from a cross-species transmission from west African chimpanzees (M and N; SIVcpz; [33, 34]) or gorillas (O and P; SIVgor; [35, 36]). HIV-2, on the other hand, describes several virus groups (A – H) that resulted from multiple cross-species transmissions from sooty mangabeys (SIVsmm; [37]). HIV-1 group M is responsible for the global pandemic and likely originated from a single transmission from chimpanzees to humans, possibly in Cameroon [33, 34], which gave rise to the current epidemic after reaching Kinshasa, located in the Democratic Republic of Congo [27]. The remaining subtypes have generally been limited to a handful of cases in Cameroon and its neighboring countries. The overall success of HIV-1 group M over its close relatives appears to be associated with various viral characteristics, as it is more pathogenic than viral strains from other groups [38].

Adaptation to host species in the previous cross-species transmissions that gave rise to HIV-1 may have allowed the virus to acquire adaptations that are better suited for replication and transmission in hominoid species [39]. SIVcpz arose from a cross-species transmission from viruses originally found in red-capped mangabeys (SIVrcm) and guenons (SIVmus/mon/gsn) [40]. In chimpanzees, SIVcpz acquired changes that allowed the virus to escape host immune factors and establish infection in hominoid species, thus resulting in one of the most devastating infectious disease outbreaks in humans.

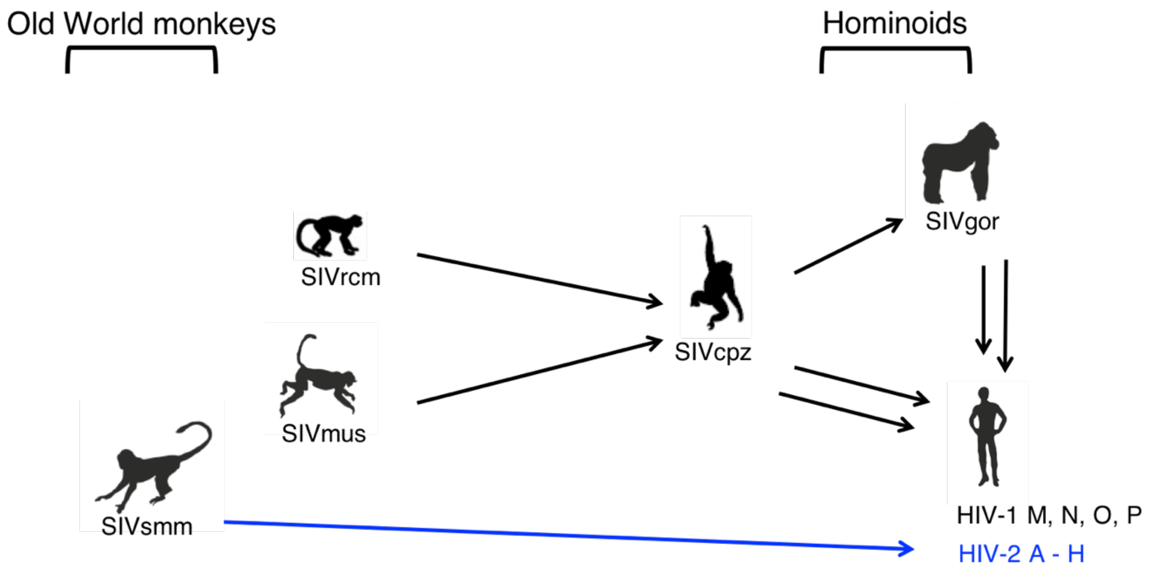

Figure 2. Origins of HIV infections from multiple cross-species transmissions of nonhuman primate SIV Transmission of SIV from Old World monkeys (left) gave rise to SIVcpz in chimpanzees, which was subsequently transmitted to gorillas (SIVgor) and humans (HIV-1, black, groups M and N). Cross-species transmission from gorillas to humans resulted in HIV-1 groups O and P. SIV strains from sooty mangabeys (SIVsmm) have also been transmitted to humans on multiple occasions, resulting in HIV-2 groups A – H (blue).

Simian immunodeficiency viruses in African primates

More than 40 African nonhuman primate species are infected with species-specific SIV strains in the wild (Figure 3) [41]. The relationships between African primates and SIVs have occurred for millions of years, supported by the observation that host proteins directly involved in the innate immune response to lentiviruses have ancient signals of virus driven evolution [42-46]. Unlike HIV in humans or SIVcpz in chimpanzees, SIV infection of natural hosts does not cause immunodeficiency despite high levels of viral replication [47-49]. This suggests that co-evolution between viruses and hosts may have allowed the virus to become less pathogenic in its natural host and progression to AIDS may be the early result of SIV host switching. Evidence has shown that natural hosts have evolved more effective means, such as disparate adaptive immune responses to non-natural hosts, to control viral infection so that it does not advance to AIDS (Figure 4; [50, 51]).

Studying SIV infection of natural hosts can provide extensive information concerning the mechanisms behind the observed lack of disease progression [52-55]. Due to their abundant availability in primate centers around the world, African green monkeys are the best studied natural SIV hosts compared to other primate species. Additionally, sooty mangabeys have been intensely studied in the context of natural SIV infection. Natural infections do share characteristics with their pathogenic counterparts, such as high viremia [47-49], high levels of immune activation [56, 57], and a significant loss of mucosal CD4+ T-cells [58, 59] during active infection. Furthermore, adaptive immune responses are unable to control virus replication throughout infection, as viral loads between natural hosts and humans are similar, demonstrating that resistance to AIDS is independent of such host defenses [60]. However, lack of progression to AIDS is tied to a number of traits that are specific to natural hosts (Figure 4). An abundance of CD4+ T-cells in the periphery of natural SIV hosts are preserved during infection and memory cells, generally thought to be part of the latent reservoir in humans, are spared from viral infection [61]. Moreover, the architecture of mucosal immune responses and lymph nodes are preserved. A lack of chronic immune activation is also observed in African green monkeys and sooty mangabeys, thus potentially protecting the animals from developing an AIDS-like illness [61]. A better understanding of the immunomodulatory adaptations seen in natural hosts will be instrumental for innovations regarding the development of anti-HIV-1 therapies. Additionally, exploring the innate immune barriers that lentiviruses must first subvert before establishing chronic infection may aid advancements in preventative measures.

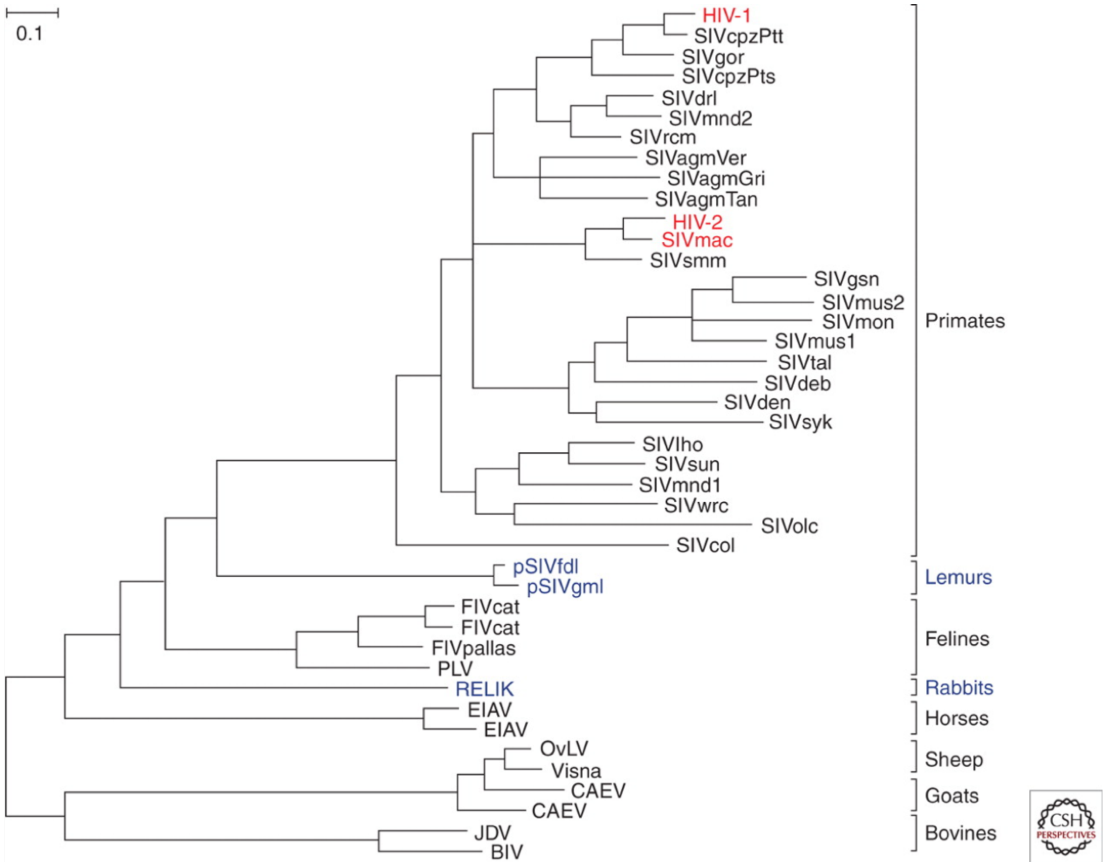

Figure 3. Phylogeny of lentiviruses
The evolutionary relationships among Pol sequences (~770 amino acids) derived from various mammalian lentiviruses; host species are indicated at the right. Exogenous viruses are depicted in black, with HIV-1, HIV-2, and SIVmac highlighted in red; endogenous viruses are shown in purple. The phylogenetic tree was estimated using maximum likelihood methods [62]. The scale bar represents 0.10 amino acid replacements per site. Reprinted with permission from [63], copyright Cold Spring Harbor Laboratory Press.

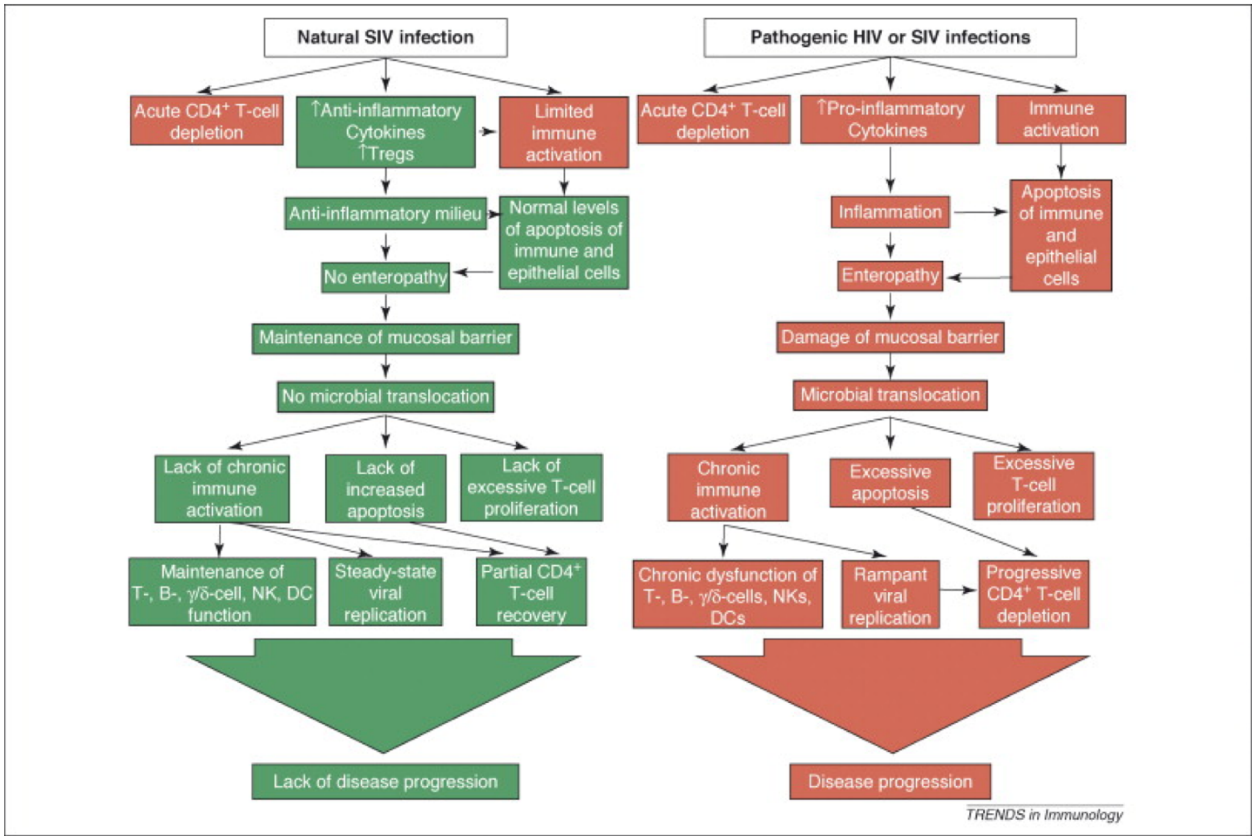

Figure 4. Comparison of events associated with natural simian immunodeficiency virus (SIV) infection and pathogenic SIV and HIV-1 infection
Differences in host response to infection are responsible for the outcome of the infection: a balanced host response in natural infection (left side) shuts down most of the factors that have deleterious consequences for the infected hosts (immune activation, apoptosis, excessive inflammation, enteropathy). In pathogenic infections, excessive host responses result in mucosal damage and consequent translocation of gut microbial products, immune system exhaustion and progression to AIDS (right side). Red boxes indicate deleterious events, whereas green boxes beneficial ones. Reprinted with permission from [53] under license number 4372760850787.

Lentiviral restriction factors

All forms of life have evolved powerful mechanisms to restrict pathogenic infections. Part of the innate immune response to lentiviral infection in mammals is comprised of a group of host proteins called restriction factors [64-67]. In general, restriction factors are constitutively expressed in host cells but they can often be upregulated by interferon signaling upon viral infection [68]. Deletion of restriction factors in mouse models generally does not have a phenotype, although an increase in susceptibility to viral infection is observed [69]. This suggests that most restriction factors only function to protect the host from viral infection. These proteins act at various steps of the viral replication cycle, including viral entry (TRIM5α), reverse transcription (APOBEC3 and SAMHD1), or viral budding (tetherin), and are thus the first-line of defense against infection (Figure 5).

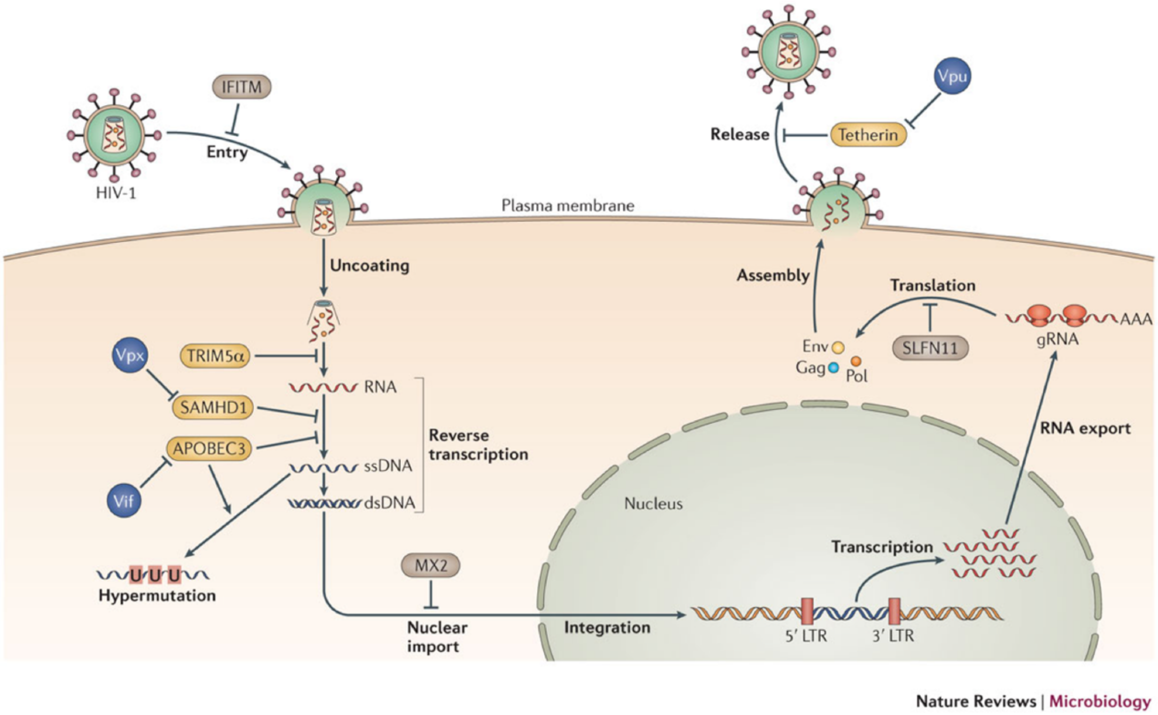

Figure 5. HIV-1 restriction and resistance factors
In the absence of virally encoded antagonists (or viral escape), host cell proteins called HIV-1 restriction factors (yellow) inhibit various stages of the replication cycle. The tripartite motif-containing protein 5α (TRIM5α) promotes accelerated fragmentation of viral cores, preventing cDNA synthesis. SAM and HD domain-containing protein 1 (SAMHD1) depletes the cellular levels of 2′-deoxynucleoside 5′-triphosphates (dNTPs), which are required for efficient cDNA synthesis. APOBEC3 (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3) proteins interfere with the processivity of HIV-1 reverse transcriptase and induce hypermutation of viral cDNA by cytidine deamination. Tetherin prevents the release of budded virions from the infected cell. Several viral proteins (blue) antagonize these cellular restriction factors. Viral infectivity factor (Vif) antagonizes APOBEC3 proteins, viral protein unique (Vpu) antagonizes tetherin, and the HIV-2 viral protein X (Vpx) antagonizes SAMHD1. HIV-1 resistance factors (brown) inhibit other stages of viral replication and are not counteracted by the virus. Myxovirus resistance 2 (MX2) prevents the nuclear import and integration of viral cDNA. Schlafen 11 (SLFN11) suppresses the translation of viral proteins. Interferon-induced transmembrane proteins (IFITMs) inhibit viral entry by interfering with membrane fusion. dsDNA, double-stranded DNA; gRNA, viral genomic RNA; LTR, long terminal repeat; ssDNA, single-stranded DNA. Reprinted with permission from [70] under license number 4372740964616.

Primate lentiviruses must adapt to evade host innate immune responses to establish infection in new hosts [65, 71]. In some cases, lentiviruses have acquired mutations in either protein targets or accessory proteins, such as *vif, vpx/vpr,* or *nef,* to counteract the host immune response (Figure 5). Here, I describe well-characterized examples of lentiviral restriction factors and viral evasion strategies.

TRIM5α is part of the tripartite motif (TRIM) containing protein family, and thus far it has been characterized as the most potently antiviral of that family [72]. It is thought to accelerate uncoating of the viral capsid, releasing the viral genome into the cell where it can be recognized by host defenses [73]. TRIM5α directly binds to retroviral capsids, therefore mutations in capsid proteins can antagonize restriction factor binding and reestablish viral infection [74, 75].

SAMHD1 (sterile alpha motif and histidine-aspartate-domain-containing protein 1) is a restriction factor capable of inhibiting the reverse transcription step of the viral replication cycle. It is a triphosphohydrolase enzyme that controls the intracellular level of deoxyribonucleoside triphosphates (dNTPs) and prevents infection of monocytes, macrophages, and dendritic cells from lentiviral infection by keeping dNTPs depleted in the cell [76]. Ultimately, when the concentration of dNTPs is too low, reverse transcription of viral RNA cannot proceed and viral infection is consequently blocked. Two lentiviral genes have evolved to antagonize the activity of SAMHD1: *vpr* and *vpx*. Primate lentiviruses encode either *vpr*, *vpx,* or both [77]. However, when both genes are present, only one restricts SAMHD1. The encoded proteins act by tagging SAMHD1 for degradation by the proteasome in a species-specific manner [78, 79].

The restriction factor tetherin anchors budding virions to the producer cell's surface, preventing their release to infect new cells. It is a transmembrane protein that is thought to physically link the cell plasma membrane to the virion envelope. Viral protein Nef antagonizes restriction by tetherin in most primate lentiviruses, by binding and removing it from plasma membranes via endocytosis [80]. In human tetherin, however, the site at which Nef would bind has been deleted [81]. Therefore, to evade host tetherin that would otherwise hinder viral infection, HIV-1 Vpu and HIV-2 Env have evolved anti-tetherin activities by promoting its degradation and sequestration in perinuclear compartments [82, 83]. Of note, only HIV-1 group M viruses encode fully functional Vpu proteins and the other subtypes have poor restriction of tetherin [84], which may explain the disparity in epidemic potential between different HIV-1 strains.

Several other examples of lentiviral restriction factors have been described in primates, such has MxB (although its relevance is controversial) [85], proteins in the IFITM family [86], and SLFN11 [87], highlighting the diversity of the innate immune response to viral infection. The existence of viral antagonists to most of these host proteins further exemplifies their impacts on viral fitness (Figure 5). Investigating the counterbalance between viruses and their hosts can be informative regarding the evolution of host immune responses.

Virus-driven evolution of lentiviral restriction factors

Host restriction factors and their viral antagonists impose selective pressures on one another to either escape (host) or preserve (virus) protein interactions [88] (Figure 6). Thus, genes encoding restriction factors often evolve rapidly. When virus-encoded antagonists are able to bind to host proteins, polymorphisms that prevent antagonism are advantageous to the host and are therefore selected in the population. In response, viruses must similarly adapt to establish the interaction and infect the host (Figure 6). This genetic conflict, or evolutionary arms race, between virus and host can be measured via amino acid substitution. A greater number of nonsynonymous mutations generally accumulate at critical binding interfaces compared to synonymous mutations. Maximum likelihood statistical methods can be used to compare the rates of nonsynonymous (dN) over synonymous (dS) substitutions and when the ratio is greater than 1, a gene or amino acid residue is said to be evolving under positive selection [88]. In primates, residues in restriction factors important for inhibiting viral infection, such as TRIM5α and SAMHD1, are under positive selection [77, 89]. Studies investigating evolutionary arms races between host and lentiviral proteins have already contributed to our knowledge regarding restriction factor function, cross-species transmissions of primate lentiviruses, and providing evidence for ancient viral infections. Further insights into virus-host interactions will help us better understand the history of lentiviral infections and ways organisms have evolved to win evolutionary battles.

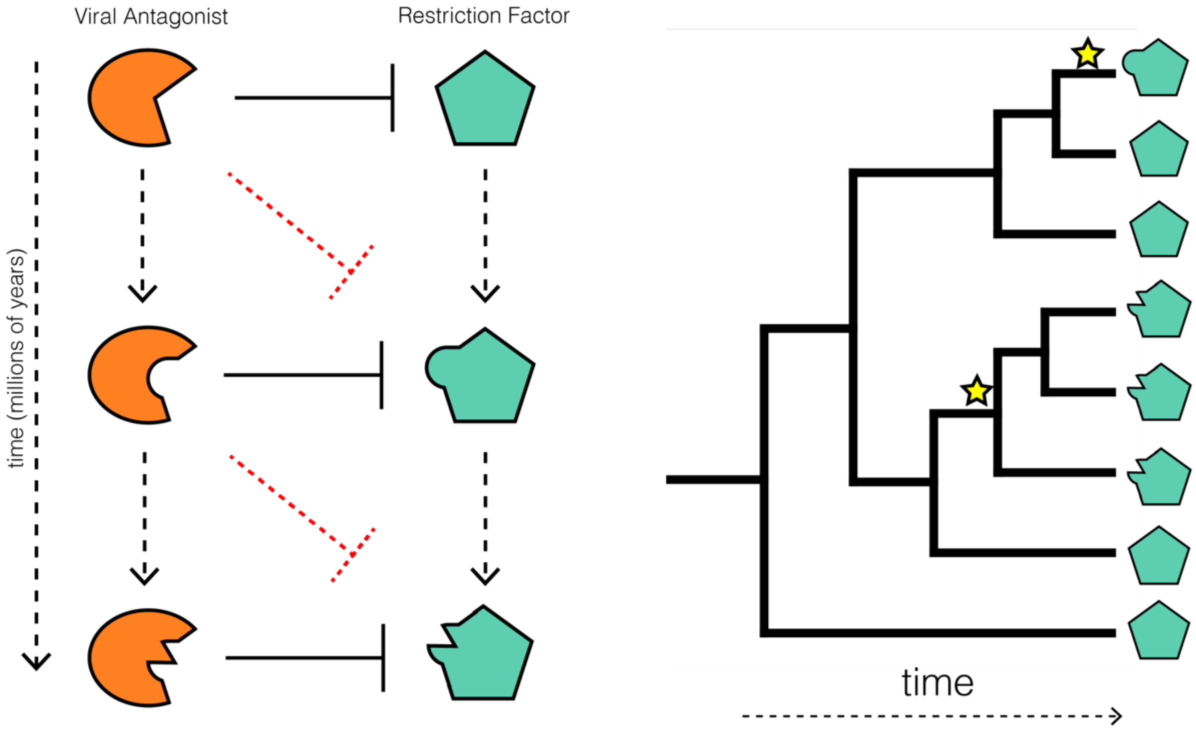

Figure 6. Genetic conflict between viral antagonists and host restriction factors
Left: The cycle of host-virus evolution in evolutionary arms races. Viral antagonists put pressure on host restriction factors to evolve to escape recognition, followed by viral adaptation to restore the interaction.
Right: The evolution of a host restriction factor constructed into a species phylogeny. The evolutionary history of virus-host interactions can be deduced by examining the age of adaptive mutations. Stars indicate mutational events.

APOBEC3 proteins

One group of restriction factors are members of the apolipoprotein B mRNA-editing enzyme catalytic-peptide like (APOBEC) family [65, 90, 91]. Genes in this family encode cytidine deaminase proteins, characterized by a signature His-X-Glu-$X_{23\text{-}28}$-Pro-Cys-$X_{2\text{-}4}$-Cys motif [92]. This motif allows for the deamination of cytidine to uridine on single-stranded DNA or RNA. The first discovered APOBEC protein, APOBEC1, edits the apo B mRNA to include a premature stop codon, which results in both full length and truncated mRNAs important for lipid metabolism [93]. These mRNAs can be differentially expressed in separate tissues, dependent on the expression of APOBEC1 [94]. Other cytidine deaminases were subsequently discovered, such as activation-induced cytidine deaminase (AID) in B-cells, which allows for modification of antibodies via somatic hypermutation and class-switch recombination [95]. However, the focus of my thesis is on the APOBEC3 group, which are critical inhibitors of retroelements and retroviral infection [90, 91]. APOBEC3 proteins were initially discovered due to a nonpermissive phenotype observed when viruses lacking the viral accessory protein Vif infected primary T cell cultures and some immortalized cell lines. However, other cell lines sustained full viral replication, called permissive cells [96, 97]. Subsequent cell fusion experiments indicated that a dominant factor was present in nonpermissive cells that could control viral replication in previously permissive cell lines [98, 99]. This factor was identified as APOBEC3G, which has since become known as one of the most potent and well-characterized APOBEC3 proteins in primates [100].

The APOBEC3 (A3) protein family has expanded and contracted over mammalian evolution through gene duplication and loss [101]. In particular, primates encode seven *A3* genes, *A3A, A3B, A3C, A3D, A3F, A3G,* and *A3H* (Figure 7), that have variable activity against retroviruses and retroelements [102]. Each contains the signature His-X-Glu-$X_{23\text{-}28}$-Pro-Cys-$X_{2\text{-}4}$-Cys motif present in all APOBEC proteins, although they are also characterized by the presence of either one (A3A, A3C, and A3H) or two (A3B, A3D, A3F, and A3G) zinc-binding domains to catalyze cytidine deamination in single-stranded DNA (Figure 7). Amino acid specificity of the zinc-binding domains classifies A3s into A3Z1, A3Z2, and A3Z3 categories [103]. Despite the differences in primary sequence, the structural features of A3 proteins are similar as each is composed of a characteristic α_1-β_1-β_2-α_2-β_3-α_3-β_4-α_4-β_5-α_5-α_6 protein fold [104].

The deamination of cytidine to uridine ultimately causes G-to-A hypermutation of its DNA substrate [105, 106]. However, proteins in the primate A3 family differ in their restriction capabilities against retroviruses and retroelements, tied to differential expression levels within specific cells or tissues and functional capabilities [102]. Inhibition of retroviruses requires packaging of A3 proteins into budding virions, where they act during the reverse transcription step of the viral lifecycle (Figure 7). A3G, for instance, is highly packaged into virions and potently restricts viral infection, while A3B is not packaged [105, 107]. Therefore, A3B has poor antiviral activity, despite strongly restricting viral infection *in vitro* [107]. Instead, A3B has been demonstrated to play an important role in restricting retroelements, like LINE-1, within primate genomes [108]. A3 proteins able to restrict retroviral infection in humans (A3G, A3D, A3F, and A3H) are constitutively expressed in T-cells and other immune cells [109]. Thus, many A3 proteins present a strong barrier to infection of the primary targets of HIV and SIV infection.

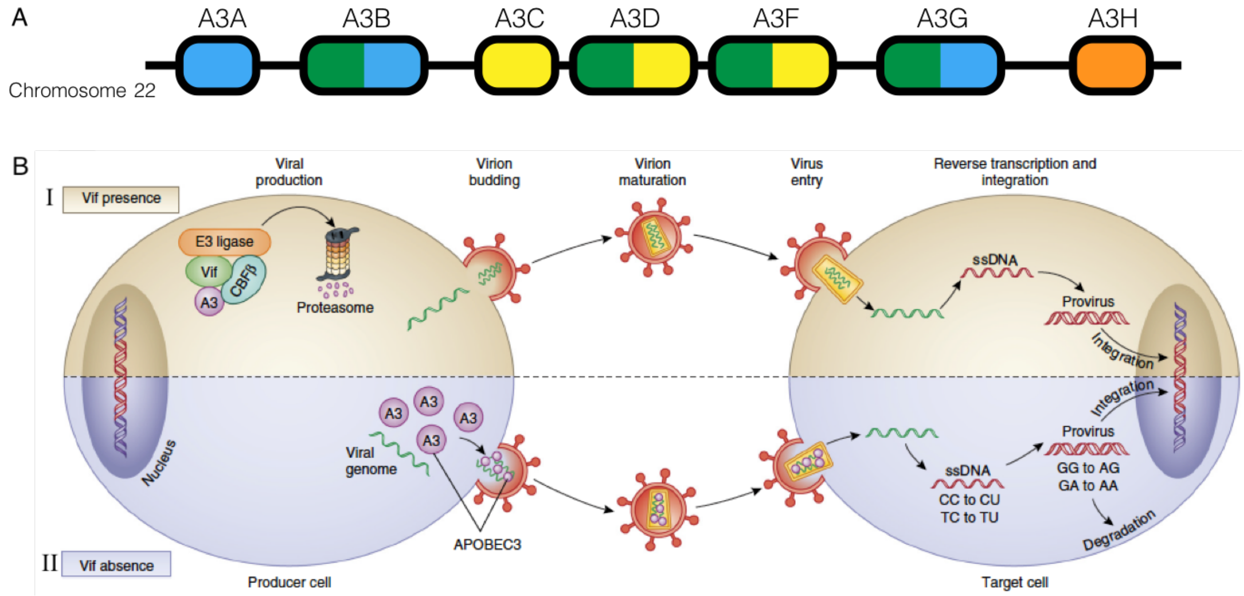

Figure 7. Antiviral activity of APOBEC3 proteins
(A) Organization of the *A3* locus on chromosome 22. Colors indicate the evolutionary history of the corresponding zinc domain according to [101].
(B) APOBEC3 proteins are degraded in the producer cell in the presence of Vif (I), but in the absence of Vif they are packaged into the budding viral particle (II). In the next cycle of infection, the APOBEC3 proteins mutagenize the viral genome during reverse transcription by deaminating cytosines to uracils in the minus-strand DNA. Mutated viral DNA may be degraded by DNA repair enzymes or integrated into the host cell genome, Reprinted with permission from [66] under license number 4372750584933.

Viral antagonism of APOBEC3 proteins

Lentiviruses have evolved to evade the antiviral effects of A3 proteins by encoding an antagonist, Vif (viral infectivity factor). *Vif* is present in almost all known lentiviruses, suggesting that it is one of the oldest accessory genes [19] . The protein hijacks a host E3 ubiquitin ligase complex composed of elongin B/C, cullin 5, Rbx2, and CBF-β to tag A3s via polyubiquitination for degradation by the proteasome, effectively decreasing the amount of A3 protein packaged into virions and rescuing the ability of the virus to infect new cell (Figure 7; [110, 111]). Viruses lacking *vif* cannot replicate in cells encoding *A3s*, demonstrating how important antagonism is for viral fitness [96, 97].

As the A3 locus in primates is so diverse, Vif must counteract an assembly of antiviral proteins [112]. Although no co-crystal structures have been resolved, A3-Vif interactions have been well-studied using mutagenesis, functional, and biochemical analyses. Electrostatic charge appears to be an important factor determining where A3 proteins and HIV-1 Vif interact (Figure 8; [104]). A3C, A3D, and A3F, for instance, share a binding interface on the α2, α3, and α4 helices, which is composed of primarily negatively charged amino acids. HIV-1 Vif, however, has a positive charge on its surface [113]. Conversely, on A3G, this binding interface has shifted to loops β1-α1, β2-α2, and β4-α4 [42, 114-116]. It is similarly thought that Vif utilizes different portions of its surface to bind various A3 proteins, although this is not completely resolved and there may be some overlap on certain surfaces [104].

The relationship between A3G and Vif has been the best characterized. It is species-specific and imposes a selective barrier preventing cross-species transmission of lentiviruses between primates [44]. Most primate lentivirus Vif proteins use amino acids 128 - 130 to antagonize A3G found in those species and both proteins have evolved in tandem, resulting in an ongoing evolutionary arms race over millions of years [42-44, 115]. Polymorphisms at these sites allow A3G to escape recognition from disparate SIV Vifs found throughout Old World monkeys in the subfamily *Cercopithecinae*, except strains circulating in species where the A3G variant is found [42]. Moreover, variation between sooty mangabeys and rhesus macaque A3G selected for SIVmac strains capable of antagonizing rhesus A3G during the cross-species transmission of SIVsmm to macaques [44]. However, Vif within the primate subfamily *Colobinae* evolved to recognize a novel insertion ^{66}SCK^{68}. While the Vif could not recognize A3G proteins with conservation at the canonical interface (128 – 130), it adapted to reestablish binding, potentially in a new region, emphasizing the importance of counteracting A3G in lentiviral evolution [42]. Exploring the evolution of other A3 proteins over extensive evolutionary history would enhance our knowledge of how Vif retains its interactions with multiple A3s and the consequences on virus-host evolution.

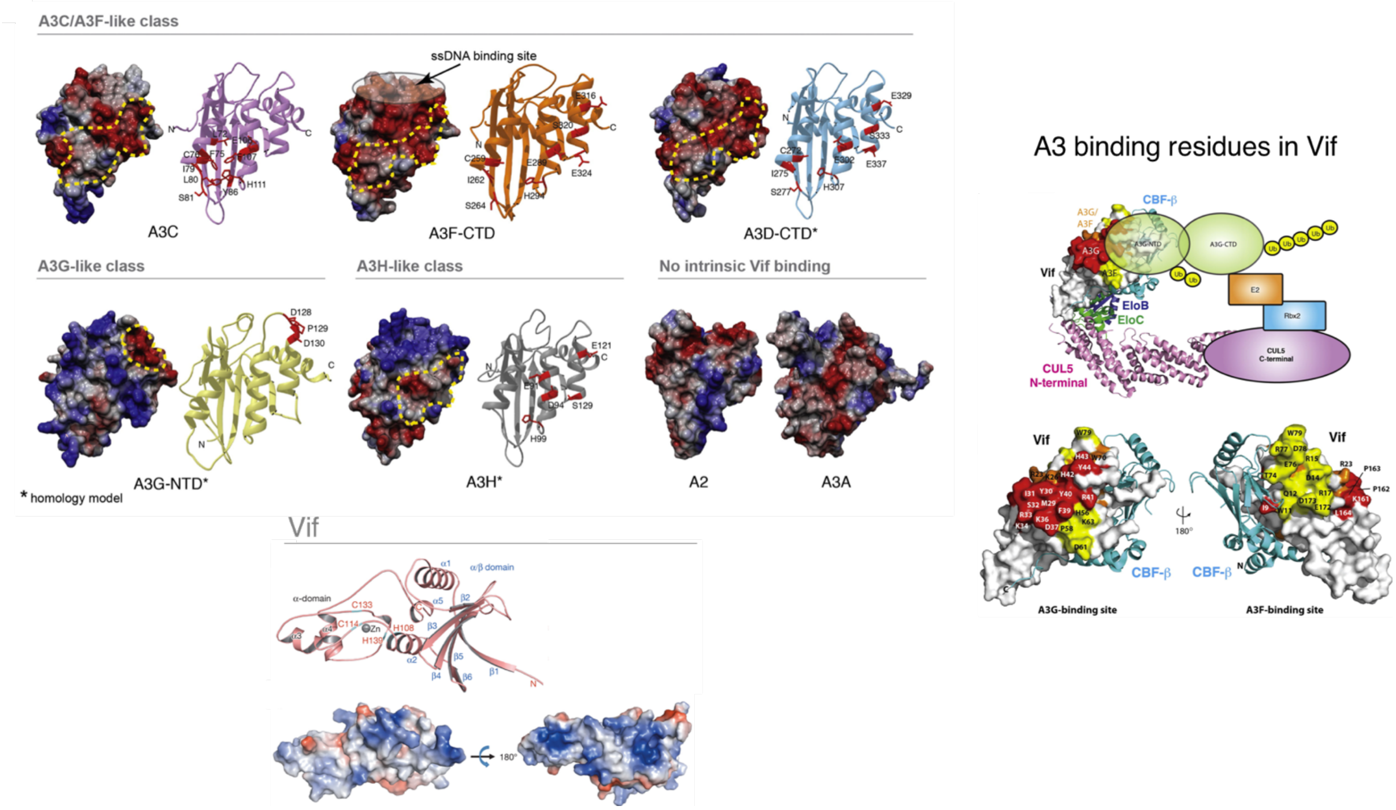

Figure 8. Structure and charge of A3 proteins and Vif

Top left: APOBEC3 proteins likely adopt three distinct types of Vif-binding sites: A3C/A3F-like, A3G-like, and A3H-like. Surface electrostatic potential and ribbon diagrams showing the three classes of A3 Vif-binding sites (yellow dashed lines) are shown. Negatively and positively charged regions are shown in red and blue, respectively. The A3C/A3F-like class is represented by A3C (PDB code: 3VOW), A3F-CTD (PDB codes: 4J4J and 4IOU) and A3D-CTD (homology model). The A3G-like and A3H-like classes are represented by homology models of the A3G-NTD and A3H, respectively. In A3C, A3F-CTD, and A3G-NTD, residues shown as red sticks in the ribbon diagrams represent experimentally determined Vif binding residues. Based on the A3F-CTD Vif-binding site, conserved residues are identified as red sticks in the A3D-CTD and A3H haplotype II homology models. Equivalent regions of the non-Vif binding A2 (PDB code: 2NYT) and A3A (PDB code: 2M65) show charge reversal and reduced hydrophobicity compared to the other A3 Vif-binding domains, thus potentially explaining its lack of Vif binding ability. Reprinted with permission from [104] under license number 4372781191471.

Bottom left: Overall structure of Vif shown in cartoon representation (ribbon). Some of the secondary structural elements of the α-domain and α/β domain of Vif are labelled. An electrostatic representation of Vif is also shown below the ribbon. Vif has a highly positively charged surface. Two views of the electrostatic surface potential map are shown. White, blue, and red indicate neutral, positive, and negative surfaces, respectively. Reprinted with permission from [113] under license number 4373740622511.

Top Right: Structure of the HIV-1 Vif-CBF-β-EloB/C-CUL5 pentameric complex (PDB code: 4N9F). CBF-β, CUL5, and EloB/C are shown in a ribbon diagram, whereas HIV-1 Vif is depicted with its molecule surface. Defined HIV-1 Vif motifs previously shown to be necessary for A3G, A3F, or A3G/A3F binding are mapped onto Vif and colored in red, yellow and orange, respectively. The putative location of the missing A3G, C-terminal domain of CUL5, Rbx2, and E2 are shown as spheres. Reprinted with permission from [104] under license number 4372781191471.

Bottom Right: Molecular surface of HIV-1 Vif illustrates two distinct binding surfaces for A3G and A3F. The A3F-binding surface is on the opposite side (180° away) from the A3G-binding interface. The presence of CBF-β on HIV-1 Vif partially restricts the A3F-binding site. The displacement of CBF-β may be necessary for the recruitment of A3F to the Vif-E3 ubiquitin ligase complex. The orientation of HIV-1 Vif in the left-hand figure is in the same orientation as in the top figure. Reprinted with permission from [104] under license number 4372781191471.

Evolution of virus-host interactions in African green monkeys

African green monkeys (AGMs) are the most numerous and widely spread species of primate found throughout sub-Saharan Africa [55]. The AGM clade is composed of 4 subspecies that are geographically distinct and infected with a species-specific strain of SIVagm: *Chlorocebus pygerythrus* (vervets, SIVagm.ver), *C. sabaeus* (sabaeus, SIVagm.sab), *C. tantalus* (tantalus, SIVagm.tan), *C. aethiops* (grivets, SIVagm.gri) (Figure 9; [117-121]). SIVagm infection in highly prevalent in AGM populations, with up to 78% of SIVagm+ adults in the wild [122]. Two other subspecies not infected with SIVagm have been described, *C. cynosuros* (malbrouck) and *C. djamdjamensis* (Bale mountains vervets), although their classification as subspecies is not universally accepted [123]. These SIVagm infected populations have been diverging for at least 4 million years, which provides a powerful system to study the evolution of the lentiviral-host arms race over relatively short evolutionary time-scales [124].

Previous studies have demonstrated that many genes involved in antiviral immunity are polymorphic in AGMs. A recent study found that a number of important genes impacting viral processes are under positive selection [125]. Indeed, multiple haplotypes of antiviral restriction factors are observed in AGMs. SAMHD1 is polymorphic in AGMs with species-specific relationships with its SIVagm viral antagonist, Vpr [78]. Haplotypes that are predominantly found in only one or two subspecies are resistant to Vpr proteins encoded by SIVagm strains from other subspecies, suggesting that SAMHD1 antagonism is an important factor in viral fitness.

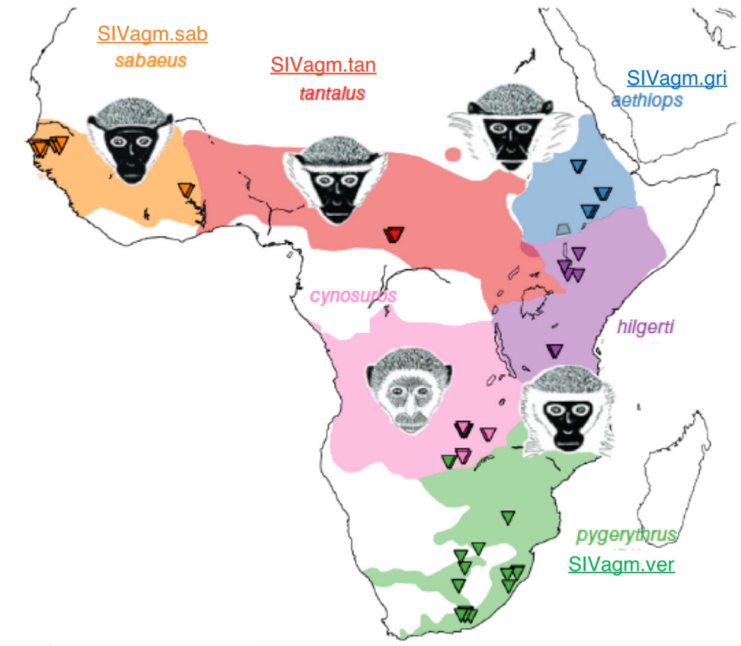

Figure 9. Distribution of AGM subspecies in Africa.

Taxon distribution and representative drawings of African green monkeys. Adapted and reprinted from [125] under license number 4372770484214.

Moreover, species-specific relationships are also observed between AGM A3G proteins and SIVagm Vif [43]. At least eight haplotypes of A3G are found in AGMs, most of which are readily antagonized by Vif proteins from all SIVagm strains. However, two haplotypes that were found only in sabaeus and grivet monkeys confer resistance to Vif-mediated degradation from other SIVagm strains. The polymorphisms responsible for maintaining antiviral activity are located in the A3G Vif-binding domain at residues 128 and 130. Haplotype VII, observed in grivets, encodes K128E and is resistant to SIVagm Vif from all strains except SIVagm.gri. Changing amino acid 128 back to a lysine (K) from glutamic acid (E) restores degradation by Vif. Similarly, haplotype VIII, found in sabaeus monkeys, encodes D130H and is only susceptible to Vif antagonism by proteins from strains in sabaeus and grivets. Making a mutation to revert the histidine

(H) at this position to aspartic acid (D) allows the protein to be degraded by Vif and rescues viral infection. This finding suggests that AGMs are currently evolving under virus-driven selective pressures.

Selection of *vif* genes that encode proteins able to counteract restriction by protective A3G haplotypes was also observed *in vivo* in this AGM cohort [43], indicating that SIVagm *vif* can readily evolve when infecting a new host expressing novel variants of A3G. When sabaeus monkeys were infected with SIVagm.ver in an experimental evolution study to better understand the role of A3G *in vivo*, *vifs* isolated from one monkey (V005) that was homozygous for AGM A3G haplotype VIII (H130/H130) were able to antagonize this A3G variant, those from individuals homozygous (D130/D130) or heterozygous (D130/H130) were not. Furthermore, the adaptation of *vifs* from V005 was only partial. Viral infection in the presence of A3G haplotype VIII was not rescued to the level of viruses encoding SIVagm.sab *vif*. This suggests that there may be other forces at play in the adaptation of SIVagm to other subspecies of AGM and perhaps other A3 proteins are influencing the evolution of *vif.*

Evidence that SAMHD1 and A3G are polymorphic and resistant to antagonism in some contexts in AGMs suggests that SIVagm continues to drive the evolution of lentiviral restriction factors and contributes to the species-specificity of viruses in AGM populations.

Loss of APOBEC3H activity in humans

Of all APOBEC3 proteins, A3H is the most evolutionarily divergent and polymorphic [101, 126]. When it was first characterized in humans, A3H was thought to have no activity against HIVΔ*vif* and related lentiviruses [127]. However, further study revealed that there are seven haplotypes of *A3H* present in human populations [128]. Haplotypes II, V, and VII encode proteins with potent antiviral activity, while haplotypes I, III, IV, and VI are less active. R105G in haplotype I and a deletion of an asparagine at position 15 (ΔN15) in haplotypes III and IV were associated with lower levels of activity. Haplotype VI encodes both of these mutations. Pulse-chase experiments revealed that inactive proteins were unstable relative to active proteins, which was also tied to changes in subcellular localization [129]. Stable A3H proteins are found primarily in the cytoplasm, while unstable proteins are also observed in the nucleus [130].

As with stability, localization to the nucleus was also linked to amino acid 105, as a G105R mutant of A3H haplotype II has a more cell-wide distribution, even in the presence of a nuclear localization signal. Using artificial protein fusions linking unstable haplotype I to stable haplotype II, it was shown that the antiviral activity of haplotype II proteins is dominant [130]. These data suggest that instability and localization are associated. Furthermore, it raised the hypothesis that stable A3H proteins are bound to a cytoplasmic host factor that prevents A3H from entering the nucleus. However, it is also possible that the observed polymorphisms between A3H haplotype I and II lead to a change in an export signal, thus changing the localization profile.

The recently resolved structure of A3H revealed that it binds to an RNA duplex as a complex of two proteins [131-133]. Consistent with previous data, treatment with RNase greatly increased its ability to mutate DNA, which indicates that binding RNA may impede deaminase activity. Loop 1, loop 7, and the α6 helix appear to be most important for binding RNA. Importantly, one study demonstrated that a lack of RNA binding was associated with a shift to the nucleus, as well as A3H antiviral activity [131]. Whether protein stability similarly relies on RNA binding has yet to be determined.

Reconstruction of ancestral A3H proteins demonstrated that the low level of antiviral activity seen in some variants of human A3H was due to a loss of function after humans diverged from the most recent common ancestor with chimpanzees (Figure 10; [129]). Along both the human and chimpanzee lineages, accumulation of nonsynonymous mutations appears to have reduced A3H antiviral activity compared to the common ancestor. However, the human A3H ancestor had a similar level of activity to the human/chimp ancestor. Nevertheless, the loss of function multiple times in the evolutionary history of humans led to the hypothesis that A3H may impart negative consequences on host fitness. Alternatively, a pressure on humans to retain A3H activity may have been lifted, rendering A3H redundant in the context of the A3 locus and genetic drift allowed for the accumulation of inactivating mutations. A broader analysis of the evolutionary history of A3H proteins in primates would reveal the overall importance of encoding a diverse A3 repertoire.

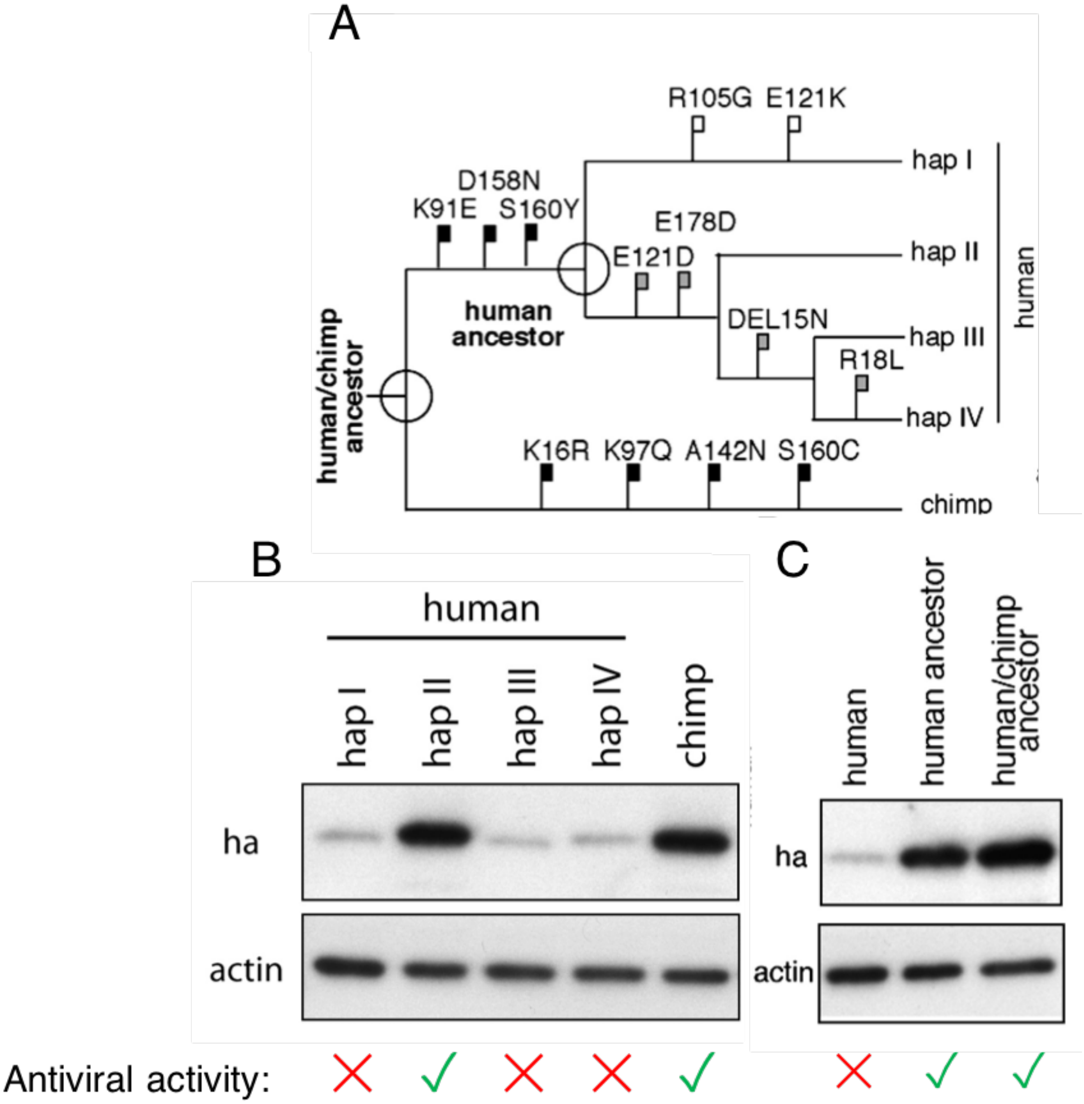

Figure 10. APOBEC3H polymorphisms in humans
The evolution of modern A3H alleles is presented as a phylogeny of the human/chimpanzee clade. Four changes (K16R, K97Q, A142N, and S160C—solid flags) occurred during the evolution of modern chimpanzees. Three changes (K91E, D158N and S160Y—solid flags) are fixed among human APOBEC3H alleles. Haplotype I accumulated two further changes (R105G and E121K—open flags) while haplotypes II–IV are the result of sequential accumulation of four additional changes (E121D and E178D for haplotypes II; plus del15N for haplotype III; plus R18L for haplotype IV—gray flags). The nodes representing the human ancestor and the human/chimpanzee ancestor are circled. (B) Western blot analysis of the four human APOBEC3H haplotypes and chimpanzee APOBEC3H. Antiviral activity is depicted below the blots. (C) Western blot analysis of human (haplotype I), human ancestor, and human/chimpanzee ancestor APOBEC3H proteins. Antiviral activity is depicted below the blots. Reprinted and adapted with permission from [129] under license number 4372761362299.

Interactions between human A3H and HIV-1 Vif

In addition to genetic polymorphisms in human A3H that impact protein stability and activity, its interactions with Vif are similarly inconsistent. Although position 121 in A3H was found to be important for the interaction with Vif (K121D makes the protein resistant to degradation) [134], the sensitivity of the protein to Vif-mediated degradation is variable between viral strains of different origin [135]. Stable A3H proteins are partially susceptible to degradation by HIV-1 LAI Vif but completely resistant to HIV-1 NL4-3 Vif [134, 136]. Additionally, primary virus strains are similarly variable in their ability to antagonize A3 proteins [137, 138]. Although A3H has been lost twice in human evolution [129], experimental evolution approaches have shown that HIV-1 Vif can evolve to antagonize A3H after accumulating changes at residues 39, 48, and 60 – 63 [138]. Vif proteins encoding V39F, N48H, and GDAK to EKGE at sites 60 – 63 are able to replicate to higher titers in PBMCs from patients encoding stable A3H. This suggests that HIV-1 truly adapts to antagonize A3H and demonstrates that active proteins are conferring a selective pressure on infecting viruses. Furthermore, stable A3H has been shown to be a barrier to cross-species transmission of SIVcpz [139]. Vifs from SIVcpz and SIVgor cannot antagonize stable A3H haplotype II and its expression in human T cell lines effectively blocked the spread of SIVcpz. Likewise, virus strains isolated from individuals with stable haplotypes of A3H are more likely to encode a Vif that can readily antagonize antiviral A3H proteins [137]. Indeed, stable A3H haplotypes are associated with a lower risk of susceptibility to infection and disease progression to AIDS in some populations [140, 141]. While infected individuals were more likely to encode unstable

A3H haplotypes in an Indian population [141], long-term non-progressors who can effectively control HIV replication without antiviral therapy more likely encoded stable A3H in a Japanese population [140].

Overall, evidence to date suggests that A3H plays a significant role in host protection against lentiviral infection. However, the evolutionary importance of A3H and its interactions with Vif in other primate species have not been fully characterized.

How has APOBEC3H evolved in African green monkeys?

The initial question my thesis project addressed was whether I could identify an evolutionary arms race in AGMs, comparable to that described for A3G-Vif [43] and SAMHD1-Vpx [78], between A3H, the most evolutionarily distinct A3, and SIVagm Vif. Identifying species-specific interactions that altered interactions with SIVagm Vif would demonstrate that A3H is an important factor for viral adaptation of lentiviruses to new primate hosts, supporting the observation that SIVcpz Vif adapted to counteract stable human A3H [139]. Another possibility is that evolution of A3H was not driven by Vif and instead reflective of a selective pressure from another pathogen, in which I would observe high levels of polymorphism in AGM A3H, but no relationship with SIVagm Vif. However, I also wanted to better understand the broader importance of A3H in primates, since its loss of activity in humans has been well-characterized. One possibility was that the evolution of *A3H* in Old World monkeys has been dynamic over evolutionary time, allowing me to more closely study the forces driving A3H activity, particularly in the context of its loss of function in some hominoid species. Alternatively, *A3H* may have been relatively conserved, potentially indicating that purifying selection might be preserving a binding interface.

My work revealed that A3H antiviral activity has been recurrently lost throughout primate evolution. Although highly polymorphic in AGMs, I observed few species-specific polymorphisms and no tested haplotypes were as potently antiviral as rhesus macaque A3H. The antiviral activity was partially dependent on protein expression level, although a mechanism independent of A3H packaging into virions similarly decreased its function. Lack of antiviral function was not associated with a robust change in localization. However, when I expanded my analysis beyond AGMs, I also found that the antiviral activity of patas monkey A3H was low. By reconstructing ancestral forms of the protein, I demonstrated that a loss of function had occurred after the most recent common ancestor of AGMs and patas monkeys, which was associated with a decrease in packaging. Conversely, rhesus macaque A3H appeared to have gained some antiviral activity from a recent common ancestor. Multiple amino acid changes were required to rescue the antiviral activity of AGM and patas monkey A3H to the level of their ancestor, suggesting that epistasis may play a role in protein function.

The fluctuation between active and less-potent proteins is not driven by Vif, which suggests that there is a cost/benefit balance conferred by active proteins. Thus, A3H does not appear to be an important component in the evolutionary arms race between A3 proteins and SIVagm Vif. Furthermore, as A3H activity has been lost in AGMs and patas monkeys, but gained in rhesus macaques, this suggests that the species barrier imposed by A3H may be variable depending on the recipient species, if it is relevant at all. My study of the evolutionary history of A3H in Old World monkeys has shown that the A3 family of proteins is dynamic and diverse, although not without the risk of a potential fitness cost.

Chapter 2. Materials and Methods

APOBEC3H cDNA Amplification and Sequencing

APOBEC3H cDNAs were cloned by nested RT-PCR (QIAGEN One-step RT PCR Kit) or PCR (Accuprime Pfx) from RNA or gDNA isolated from AGM peripheral blood mononuclear cells or cell lines. Sample origins and extractions have been previously described [43, 78]. For each sample, PCR products were amplified and sequenced using primers designed to amplify African green monkey A3H (Forward: CACGAATTCGCCACCATGTATCCATACGATGTTCCAGATTACGC TGCTCTGCTAACAGCCAAA Reverse: CACGAGCTCATCTTGAGTTGAGTGT). Primers for gDNA amplification were designed to target intronic regions. Heterozygous sequences were cloned using pGEM T-Easy vector system (Promega) and TOPO TA cloning kit (Invitrogen) to analyze individual clones. Additional primate sequences were obtained from gDNA isolated from immortalized cell lines using QIAGEN DNeasy Blood & Tissue Kit. The following cell lines from Coriell Cell Repositories (Camden, NJ) were used: patas monkey (Erythrocebus patas; ID no. 6254), De Brazza’s monkey (Cercopithecus neglectus; PR01144), Wolf’s guenon (Cercopithecus wolfi; PR01241), mustached guenon (Cercopithecus cephus, PR00527), Allen’s swamp monkey (Allenopithecus nigroviridis; PR00198), and Francois’ Leaf monkey (Trachypithecus francoisi; PR01099). Talapoin (Miopithecus talapoin; OR755) cells were obtained from Frozen Zoo (San Diego, CA).

Expression Constructs and Plasmids

Primate A3H genes were cloned from cDNA. A 5’ hemagluttinin (HA) tag was added via PCR (Forward: GTGGTGGAATTCATGTATCCATACGATGTTCCAGATTACGCTGCTCTGCT Reverse: CTAGACTCGAGTCATCTTGAGTT). The products were digested using EcoRI/XhoI restriction enzymes and ligated into a mammalian pcDNA 3.1 vector (Invitrogen, #V79020). Site-directed mutagenesis was completed with the QuikChange II Site-Directed Mutagenesis Kit (Agilent, #200524) to construct all ancestral genes and mutants. The A3H gene from patas monkeys was generated by gene synthesis (IDT) and cloned into the pcDNA 3.1 backbone. A3H genes from AGM and rhesus macaque were codon-optimized based on usage frequencies in primates (human and rhesus macaque) to remove rare codons within the gene in Geneious (Biomatters Ltd.). Codon-optimized sequences were generated by gene synthesis and cloned into pcDNA 3.1.

Cell Lines, Transfections, and Western Blot Analysis

HEK293T, HeLa, and Cos7 cell lines (ATCC) were maintained in Dulbecco’s modified Eagle’s medium (DMEM) with 10% fetal bovine serum (Corning, #35-015-CV) and 100 μg/mL penicillin/streptomycin (Gibco, #15140-122) at 37°C. SupT1 cells (ATCC) were maintained in RPMI 1640 with 10% fetal bovine serum and 100 μg/mL penicillin/streptomycin in the same conditions. Transfections were done in serum-free DMEM with TransIT-LT1 transfection reagent (Mirus Bio, #MIR 2305) at a reagent:plasmid DNA ratio of 3:1. For western blot analysis, cells were lysed in ice-cold

NP40 buffer (0.5% NP40, 20mM NaCl, 50mM Tris pH 7.5) with protease inhibitors (Roche Complete Mini, EDTA-free tablets, #11836170001). Lysates were quantified using a Pierce BCA Protein Assay Kit (Thermo Scientific, #23225) and 10 μg of protein was resolved by SDS-PAGE, transferred to a PVDF membrane, and probed with anti-HA (BioLegend, #901503) and anti-actin (Sigma, #A2066) or anti-vinculin (Proteintech, #66305-1) antibodies at a 1:2000 dilution. Samples from viral supernatants were probed with a p24 antibody (NIH-ARP, #3537) at a 1:1000 dilution. Anti-mouse or anti-rabbit secondary antibodies were used at a 1:5000 dilution (Santa Cruz Biotechnology, sc-2005, sc-2004).

Immunofluorescence

HeLa cells were seeded onto 18mm (VWR, #48380 046) coverslips seeded with 4 x 10^4 cells/mL and transfected with 500 ng of A3H-expressing plasmids the next day. 48 hours after transfection, the coverslips were fixed in 2% paraformaldehyde, permeabilized in 0.5% PBS/Triton-X, and blocked in PBS/BGS. HA-tagged proteins were detected using the same HA antibody used for western blots at a 1:1000 dilution followed by an anti-mouse AF488 antibody at 1:400 (Invitrogen, #A11001). Nuclei were stained in SlowFade Gold antifade reagent with DAPI mounting media (Life Technologies, #S36939). Images were taken on a Nikon E800 microscope.

Phylogenetic Analysis

A3H AGM sequences were analyzed phylogenetically using a Bayesian Monte Carlo Markov chain (MCMC) approach implemented in BEAST v1.7.1. Sequence alignments were constructed using MAFFT align function in Geneious (Biomatters Ltd.) and underwent 10,000,000 MCMC generations using HKY85 substitution model, gamma site heterogeneity model, estimated base frequencies, and constant population size coalescent as the tree prior.

Positive Selection Analysis

The 27 primate A3H sequence alignment was analyzed using HyPhy GARD analysis to ensure there was no recombination in the gene [142]. The species phylogeny [124] was input into the CODEML sites model of PAML [143] along with the nucleotide alignment to detect positive selection at individual sites. The p-value was calculated by twice the difference in log-likelihood between models M7 and M8 as well as M8 and M8a with two degrees of freedom. Analysis was conducted with both the F3x4 and F61 codon frequency models with omega values of 0.4 and 1.5. Data from F3x4 and F61 models are shown in Table 1. Positively selected sites were categorized as those with an M8 Bayes Empirical Bayes posterior probability greater than 98%.

Ancestral Reconstruction

The ancestors at specific nodes within the Old World monkey clade were reconstructed using the FASTML webserver (fastml.tau.ac.il; last accessed March 2018; [144]). The 27 primate A3H sequence alignment was used in conjunction with the species tree to generate marginal reconstructions of codon sequences.

Single-Cycle Infectivity Assays

HEK293T cells were plated in 1mL in 12-well plates at 1.25 x 10^5 cells/mL. After cells reached between 50 – 70% confluency, they were co-transfected with 250 ng of A3H or empty expression plasmid, 600 ng of a luciferase-encoding proviral plasmid with a frameshift mutation in env, and deletion of vif, and 100 ng of L-VSV-G (vesicular stomatitis virus glycoprotein) for pseudotyping in 100 μL serum-free medium with TransIT-LT1 transfection reagent (Mirus Bio). Supernatants containing virus were harvested after 48 hours and clarified through 0.2-micron filters. Viral titers were determined by measuring reverse-transcriptase (RT) activity by qPCR as described previously [145]. In short, viral supernatants were lysed in 2X lysis buffer (0.25% Triton X-100, 50 mM KCl, 100 mM Tris-HCl, 40% glycerol) in the presence of 4U RNase inhibitor (Fermentas, #EO0382). qRT-PCR reactions were set up with an MS2 RNA template using the Takyon Rox SYBR MasterMix dTTP Blue kit (Eurogentec, #UF-RSMT-B0101) alongside a standard curve made with a stock virus of previously determined titers. The primers used to amplify duplicate reactions were: TCCTGCTCAACTTCCTGTCGAG (forward) and CACAGGTCAAACCTCCTAGGAATG (reverse). qRT-PCR was performed on an ABI QuantStuido5 Real Time PCR machine. 2000 mU/mL was used to infect SupT1 cells plated at 2 x 10^4 cells/well in a 96-well plate in media supplemented with 20 μg/mL DEAE-Dextran. Infections were done in triplicate for 48 – 64 hours. Luciferase activity was measured with Bright-Glo Luciferase Assay Reagent (Promega, E2620) on a LUMIstar Omega luminometer.

Packaging Assay

Packaging of A3 proteins into virions was measured by co-transfection of 600 ng of A3 plasmid with 1000 ng of the proviral plasmid described for single-cycle assays into a 6-well plate seeded with HEK293T cells at 2 x 10^5 cells/mL. Two days after transfection, cell lysates were collected and 1.5 mL of supernatant was clarified through 0.2-micron filters. Viral supernatants were concentrated by pelleting by centrifugation at 13,000 RPM for 60 minutes. Media was aspirated from the virion pellet until ~10 uL remained and lysed with 4x NuPAGE LDS Sample Buffer (Invitrogen, NP0007). Virion and cell lysates were analyzed by western blot.

Statistical Analysis

Data analyses were done using GraphPad Prism 7.0 software. All data are shown as mean ± SEM. Statistical analysis was performed using unpaired t-tests with two tailed, 95% confidence. P values less than 0.05 were considered statistically significant.

Accession Numbers

The GenBank accession numbers for new Old World monkey A3H sequences, including De Brazza’s, Wolf’s Guenon, mustached guenon, patas monkey, talapoin, Francois’ Leaf monkey, Allen’s swamp monkey, and AGM haplotype, reported here are MH231602 – MH231609.

Chapter 3. Recurrent loss of APOBEC3H activity during primate evolution

Introduction

The seven members of the APOBEC3 (A3) gene family in primates encode cytidine deaminases involved in innate immune defense against retroviruses and retroelements [90, 91, 146]. Four A3 enzymes are known to potently restrict the replication of lentiviruses like simian immunodeficiency virus (SIV) and human immunodeficiency virus (HIV): A3D, A3F, A3G, and A3H [102]. A3 proteins are packaged into budding virions and cause G-to-A hypermutation of viral DNA, although deamination-independent modes of restriction have also been characterized [147, 148]. Hypermutation of viral DNA results in detrimental mutations that render the virus inactive and thus protects new cells from infection. However, lentiviruses have evolved a mechanism to evade A3 restriction by encoding a viral antagonist, Vif, which binds and targets A3 proteins for proteasomal degradation via a cellular E3 ubiquitin ligase complex [66]. Given that A3-Vif interactions often act in a species-specific manner, adaptation of Vif to host A3 proteins is important for successful adaptation of lentiviruses to their hosts [42, 43, 45, 139, 149].

Various A3 proteins differ in their ability to restrict viral infection. For instance, A3G is the most potent A3-mediated inhibitor of HIV-1, while A3A and A3B have limited antiviral potential [102, 107]. In humans, the A3H protein is especially remarkable because multiple polymorphisms drastically impact its antiviral activity [127-129, 150].

Two independent mutations have occurred in human evolution that destabilized A3H [129]; haplotypes that encode R105G or a deletion of amino acid 15 make unstable proteins (haplotypes I, III, IV, VI) that have lost antiviral activity, while those without these changes make stable proteins (haplotypes II, V, VII) that potently restrict HIV-1 [127-129, 138]. Stability and antiviral activity of human A3H has been further linked to subcellular localization: unstable/poorly-antiviral proteins are more nuclear, while stable/antiviral proteins remain cytoplasmic [130]. A3H haplotype I has been associated with breast and lung cancer [151], as has another nuclear A3, A3B [152], suggesting that in some cases A3 activity may be detrimental for the host. While such events have occurred in humans, examples of gains or losses of A3 activity over evolutionary time in other primates have been less explored.

The A3H genetic polymorphisms present in human populations also impact the interactions between human A3H and HIV-1 Vif [134, 136-138]. Stable A3H proteins are only partially susceptible to degradation by Vif from the LAI isolate of HIV-1 and not at all by HIV-1 NL4-3 Vif [134, 136]. However, studies using human cohorts encoding different haplotypes of A3H have shown that Vif from some primary virus strains are able to antagonize stable A3H proteins [137, 138, 140, 141]. These studies suggested that there is selection *in vivo* for Vif strains that counteract the stably expressed forms of A3H in infected people [137, 138, 140, 141]. Furthermore, the cross-species transmissions that led to adaptation of SIV from monkeys to chimpanzees to humans, giving rise to HIV-1, involved adaptation of Vif to antagonize the A3 proteins found in each host [42, 43, 139, 149], including SIVcpz Vif to antagonize human A3H [139].

The study of the evolution of lentiviral-host interactions within Old World monkeys has provided insights into the longer-term dynamics of the evolutionary arms race between host antiviral proteins and their lentiviral targets [42, 45, 79]. African green monkeys (AGMs), in particular, provide a unique opportunity to assess the evolutionary forces governing interactions between lentiviruses and their hosts since the genus Chlorocebus encompasses four geographically distinct subspecies (vervets, sabaeus, tantalus, and grivets) each of which are infected with species-specific subtypes of SIVagm [41, 124, 153]. These SIV subtypes are named by the subspecies they infect: SIVagm.ver, SIVagm.sab, SIVagm.tan, and SIVagm.gri [41]. Furthermore, this system is particularly powerful because although these subspecies are closely genetically related, enough divergence exists that has allowed species-specific lentiviral infections to occur.

Previous studies have demonstrated that many genes involved in antiviral immunity are polymorphic in AGMs and some changes at the protein level are critical for interactions with viral antagonists [43, 78, 125]. A3G in particular was discovered to encode species-specific polymorphisms [43]. Amino acid changes in the Vif binding domain of AGM A3G (sites 128 and 130) confer protection against SIVagm strains from other subspecies. For example, K128E, found in grivet monkeys, is resistant to all Vifs except SIVagm.gri, while D130H, found in sabaeus monkeys, is resistant to Vifs from SIVagm.ver and SIVagm.tan [43]. This demonstrates that Vif continues to drive the evolution of A3G and contributes to the species-specificity of lentiviruses in AGM populations. It is unknown whether A3H has a similar role in these primates.

In this study, we asked whether we could identify a host-virus “arms-race” between A3H in AGMs and Vif proteins encoded by the SIVs that infect these species. Surprisingly, we found that although A3H is highly polymorphic in AGMs, the antiviral activity of A3H has largely been lost. The reduced antiviral activity is in part caused by lower protein expression levels, although amino acid changes also independently lower antiviral activity. By reconstructing ancestral A3H proteins spanning evolution in AGMs and in other Old World monkeys, we find that there has been recurrent loss of A3H activity in some, but not all, Old World monkeys. While higher expression levels generally increase viral inhibition, we also identified amino acids that affect A3H restriction, which map to regions implicated in RNA binding [131-133]. Thus, our data support a model where A3H antiviral activity has been repeatedly lost throughout evolution. This argues that there is a longer scale dynamic between the cost and benefit for A3H function in primates that is not necessarily driven by interactions with its antagonist.

Results

A3H is highly polymorphic in AGM subspecies, but all tested alleles have low antiviral activity

Polymorphisms in human A3H are known to affect its antiviral activity as well as its interactions with Vif [128, 129, 136-138]. Similarly, polymorphisms in A3G in AGMs impact interactions with Vif, suggesting an ongoing and ancient genetic conflict between A3G and SIV Vif in AGMs [43]. These observations prompted us to explore the genetic landscape of A3H in African green monkeys and other Old World monkeys to determine if A3H evolution has been driven by Vif over a broad evolutionary scale. First, we sequenced A3H in 50 AGM samples collected from all four subspecies infected with a species-specific SIV: vervet, sabaeus, tantalus, and grivet monkeys. The mitochondrial DNA of these animals had also been previously sequenced and was confirmed to cluster by AGM subspecies [43]. Sequence analysis of 80 independent A3H genes revealed 34 single-nucleotide polymorphisms (SNPs), 6 of which are synonymous changes at a frequency higher than 5%. 28 SNPs are nonsynonymous; 23 are found in more than one individual (Table 1). Most polymorphisms are represented in all four subspecies (Figure 11 A and B), except for low frequency variants only found in one AGM and two nonsynonymous mutations found in two vervet sequences at amino acid residues 113 and 116. Phylogenetic analysis showed that A3H sequences from all AGM subspecies are paraphyletic (Figure 11A), similar to a previous study examining A3G in these species [43].

Table 1. Polymorphisms identified in AGM *A3H* sequences
[*] Amino acid residues only identified in one animal
[a] Location in AGM A3H modeled onto a previously described pig-tailed macaque A3H ([133], PDB 5W3V)

Site	Observed amino acids	Predicted location in structure[s]
17	R/H	Loop 1
18	Y/R/H/L*	
20	S/N	
25	P/R	
41	T/M*	Loop 2
44	R/K*	β-sheet
46	H/Q	
53	H/D	Loop 3
73	C/S*	Loop 4
75	R/Q	β-sheet
113	Y/C	Loop 7
116	R/H	
117	R/P*	
127	C/S	α4
128	G/R	
130	R/Q	Loop 8
134	E/K	β-sheet
138	L/R	α5
139	P/Q	
152	K/E	Loop 10
164	D/E	α6
171	Q/R	
181	K/E	
195	N/S	Not resolved in structure
204	S/A	

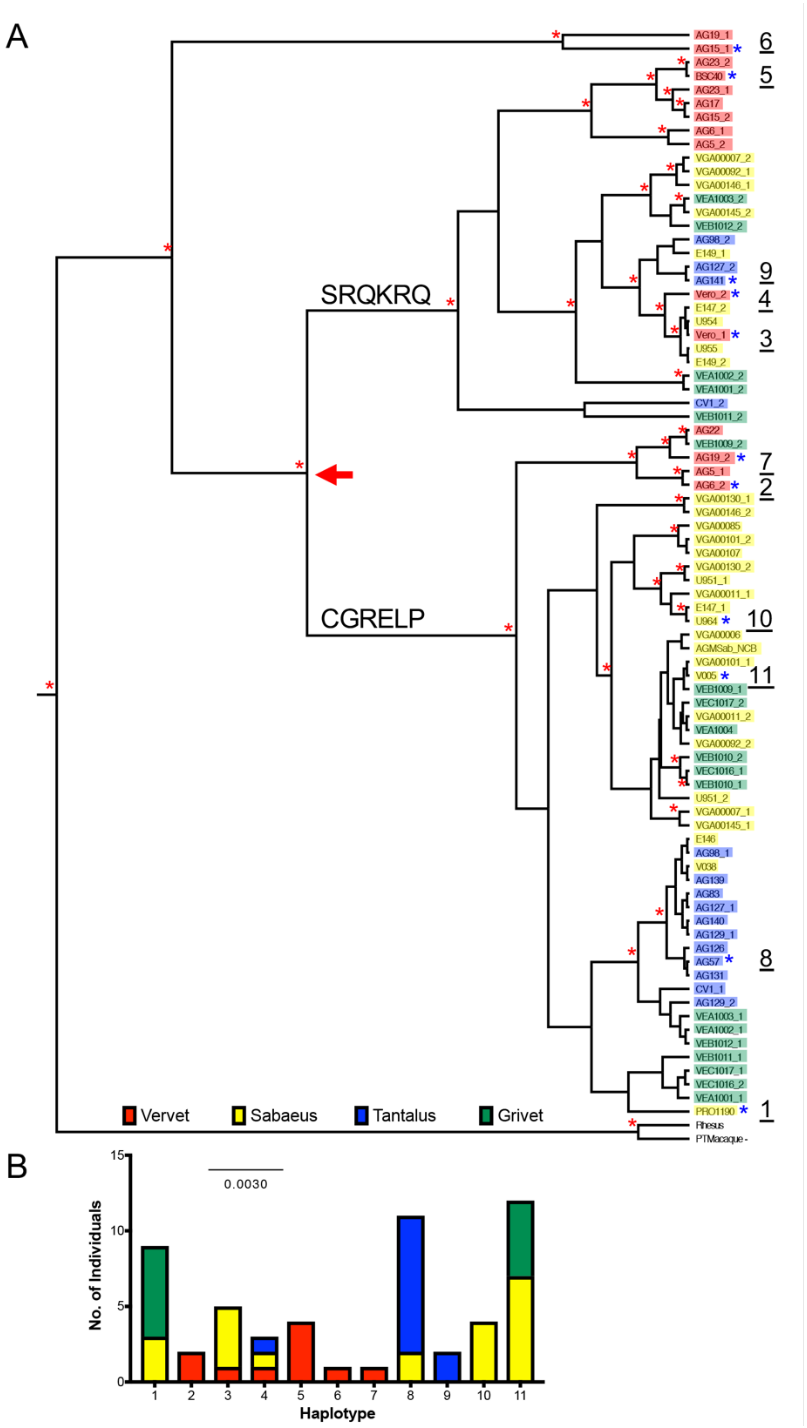

Figure 11. Sequence and Phylogenetic Analysis of *A3H* in African green monkeys
A. The evolutionary relationship between 80 full-length AGM *A3H* genes was inferred by Bayesian MCMC phylogenetic reconstruction. Red asterisks (*) show nodes that have a posterior probability > 0.5. Colored boxes demonstrate the subspecies of origin (red = Vervet, yellow = Sabaeus, blue = Tantalus, and green = Grivet). Blue asterisks (*) denote cloned haplotypes, numbered 1 – 11. The red arrow indicates divergence between two clades. **B.** Number of AGM individuals encoding haplotypes 1 – 11, color coded by subspecies similar to *A3H* phylogeny.

Nonsynonymous SNPs are spread throughout the protein (Table 1), although one group is clustered between amino acids 127 – 139. The changes are tightly linked and result in divergence of the phylogenetic tree between two clades (Figure 11A, red arrow). Haplotypes with a CGRELP motif compose one clade on the tree, while the other has an SRQKRQ motif. Residues 127 and 128 are located on the α4 helix, which has been implicated in A3H-Vif binding [154, 155]. Furthermore, the regions of A3H with the greatest number of nonsynonymous mutations are in the predicted loops 1 and 7, and the α6 helix (Table 1), which were recently shown by structural studies to be involved in an interaction between A3H and a co-crystalized RNA [131, 133]. These results indicate that, similar to A3G, SAMHD1, and other antiviral genes in AGMs [42, 78, 125], there is extensive polymorphism in AGM A3H.

The presence of numerous polymorphisms suggests there may be functional consequences for either antiviral activity or Vif antagonism. To determine whether or not nonsynonymous polymorphisms in A3H impact antiviral activity, we tested the ability of A3H protein variants to restrict lentiviruses. We cloned haplotypes, numbered 1 – 11 and representative of each clade in the phylogenetic tree, into a mammalian expression vector for functional analysis (Figure 11B). These haplotypes also represent all four subspecies of AGMs. Out of the 28 total nonsynonymous SNPs found in AGMs (Table 1), 11 unique protein sequences were tested and 20 polymorphic sites were characterized (Table 2).

Table 2. Polymorphic sites in tested A3H haplotypes

[a]Haplotype number of A3H allele

[b]Amino acid position with nonsynonymous mutations across the tested haplotypes. Boxes colored red have a different amino acid residue relative to AGM A3H haplotype 1.

	Amino acid																		
Haplotype[a]	17[b]	18	20	25	53	75	113	116	127	128	130	134	138	139	152	164	171	182	195	204
1	R	H	S	P	D	Q	Y	R	C	G	R	E	L	P	K	Q	Q	K	S	S
2	R	H	S	P	D	Q	Y	R	C	G	R	E	L	P	K	Q	Q	K	N	A
3	H	H	S	R	H	R	Y	R	S	R	Q	K	R	Q	E	R	R	E	S	S
4	R	H	S	P	H	R	Y	R	S	R	Q	K	R	Q	E	R	R	E	S	S
5	R	H	S	P	D	Q	Y	R	S	R	Q	K	R	Q	E	R	R	E	N	A
6	R	R	N	R	D	Q	C	H	C	G	Q	K	L	Q	E	R	R	E	S	S
7	H	H	N	R	D	Q	Y	R	C	G	R	E	L	P	K	Q	Q	K	N	A
8	R	H	N	P	D	Q	Y	R	C	G	R	E	L	P	K	Q	Q	K	S	S
9	R	H	S	P	H	R	Y	R	S	R	Q	K	R	Q	E	R	R	E	N	A
10	R	H	S	R	D	Q	Y	R	C	G	R	E	L	P	K	Q	Q	K	S	S
11	H	H	S	R	D	Q	Y	R	C	G	R	E	L	P	K	Q	Q	K	S	S

A3H expressing plasmids were co-transfected into HEK293T cells for single-round infectivity experiments with HIVΔenvΔvif or SIVagm.TANΔenvΔvif proviruses and a VSV-G expression plasmid for pseudotyping to measure antiviral activity. Viral supernatants were collected, normalized for RT activity, used to infect SupT1 cells, and luciferase activity from a gene encoded in the proviral plasmid was measured relative to a no A3 control (see Methods). We found that none of the AGM A3H variants restricted lentiviruses as potently as the A3H from rhesus macaque (Figure 12A), an Old World monkey that has been previously characterized for its A3H activity [102, 127, 129]. Rhesus A3H restricted viral infection of HIVΔvif by approximately 21-fold, while AGM A3H variants restrict viral infection by no more than 3-fold, and some not at all (Figure 12A). Although activity against HIV-1Δvif correlates with activity against other lentiviruses, we also validated this result using SIVagm.TANΔvif, a strain originally isolated from tantalus monkeys. We found that the AGM A3H haplotypes were also poorly restrictive against SIVagm.TANΔvif compared to the activity of rhesus macaque A3H (Figure 12A). SIVagm.TANΔvif was observed to be more sensitive to restriction by all A3H proteins tested, similar to a previous study from our lab using A3G [43], although the difference in viral inhibition for specificity between inhibition of HIV-1 or SIVagm.TAN within haplotypes was not statistically significant (unpaired t-test, $p > 0.05$). Additionally, to ensure that the epitope tag was not interfering with protein function, we removed the HA tag and compared the antiviral activity of tagged versus untagged proteins. We found that, relative to rhesus macaque A3H, there are no significant differences between the restrictive capabilities of tagged and untagged proteins (Figure 13, unpaired t-test, $p>0.05$). This indicates that A3H variants encoded by AGMs have poor antiviral activity against at least two separate lentiviruses and is not due to species-specificity (Figure 12B).

The differences in restriction could be explained by changes in expression since we found that no variant of AGM A3H is as strongly expressed as rhesus macaque A3H in HEK293T cells (Figure 12B, top). This lower expression of AGM A3H proteins relative to rhesus macaque A3H was not due to the species of origin of the cells used for transfection since when we transfected AGM-derived Cos7 cells, the protein expression levels of AGM A3H were similarly poor relative to the expression of the rhesus A3H (Figure 12B, bottom). These data indicate that lower protein expression levels are correlated with less potent antiviral activity, similar to unstable haplotypes of human A3H.

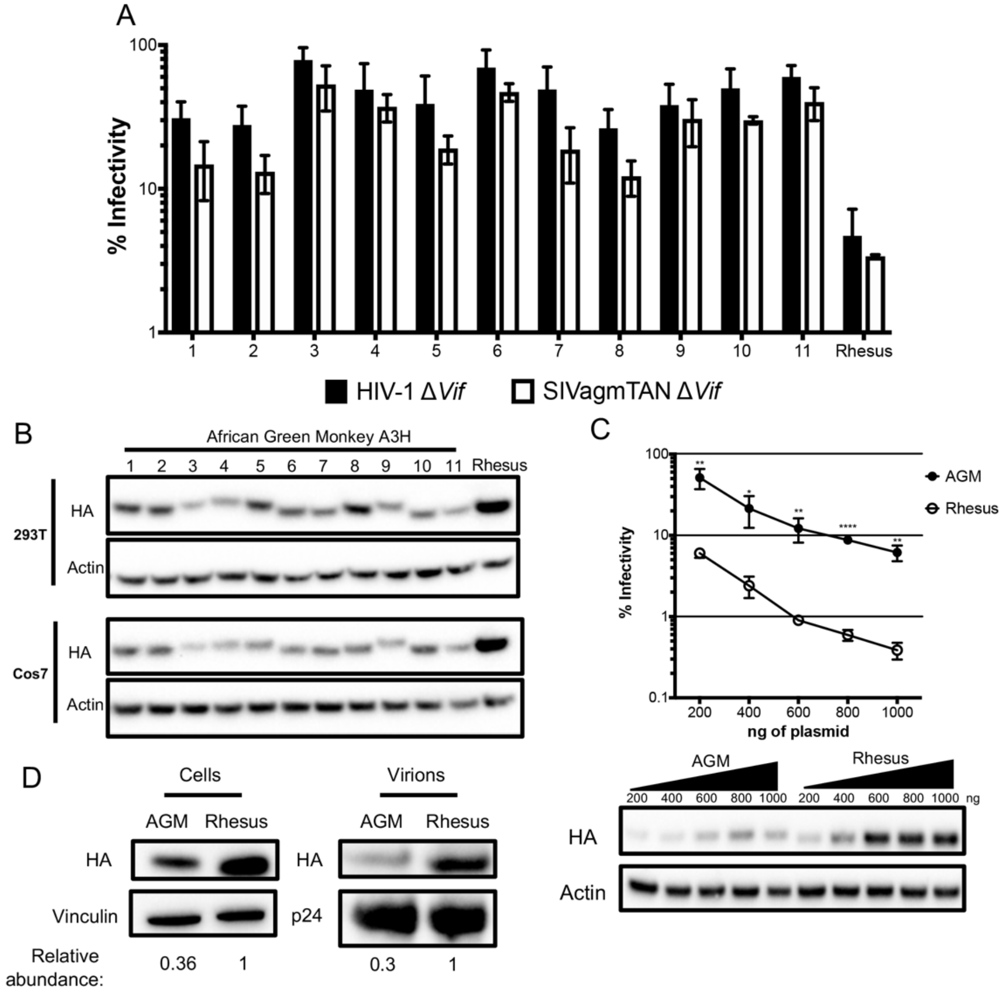

Figure 12. Antiviral activity of A3H is lower in AGMs than in another Old World monkey
A. Single-cycle infectivity assays were performed in the presence or absence of A3 proteins against HIVΔ*vif* (black) and SIVagmΔ*vif* (white). Rhesus macaque was included as a positive control. Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. All samples were statistically significantly different than rhesus macaque A3H, except AGM haplotype 7 against SIVagm.TAN ($p = 0.0502$) as measured by unpaired t-tests. Furthermore, no significant differences in restriction between HIV and SIVagm.TAN within individual haplotypes were found. **B.** Western blot analysis of HA-tagged AGM A3H protein expression in human (HEK293T) and AGM (Cos7) cell lines. The different size bands for different AGM A3H haplotypes is reproducible. β-Actin is shown as a loading control. **C.** Top: Single-cycle infectivity assay of HIVΔ*vif* in the presence of increasing amounts of A3H-expressing plasmids. AGM A3H haplotype 1 (black circles) and rhesus macaque A3H (open circles) are compared. Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. Statistical differences were determined by unpaired t-tests: $^*p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, and $^{****}p \leq 0.0001$. Bottom: Western blot analysis of protein expression levels with the same amounts of plasmid in the top panel. β-Actin is shown as a loading control. **D.** Packaging of A3H proteins into virions analyzed by western blot. Relative abundance in the cellular expression (left) and virion incorporation (right) were determined compared to rhesus macaque A3H.

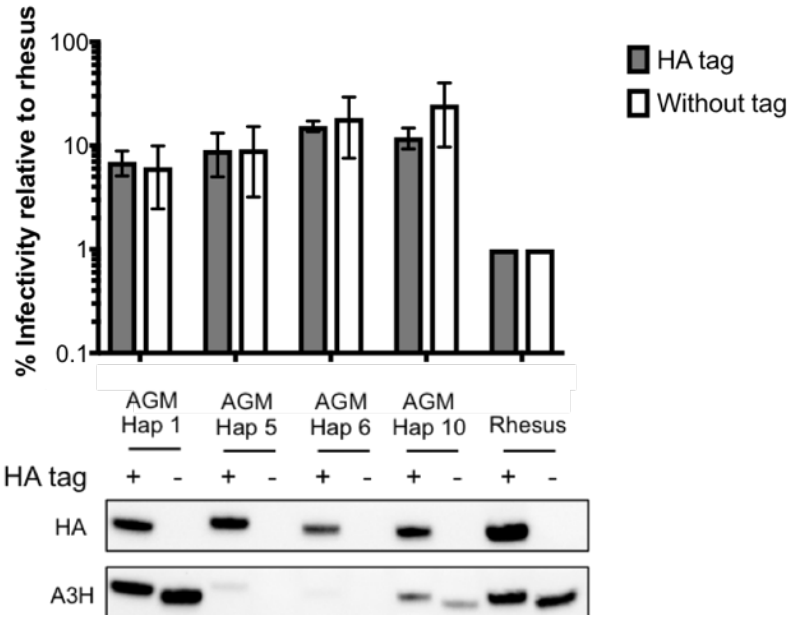

Figure 13. AGM A3H proteins without an epitope tag still have lower antiviral activity than rhesus A3H Single cycle assays were performed using constructs described in the manuscript (with an HA tag, gray bars) and constructs without the epitope tag (white bars). The relative infectivity of each protein variant compared to the infectivity of rhesus A3H in individual experiments was determined and the average of three replicate experiments, each with triplicate infections (± SEM) was graphed. Below the graph is the corresponding western blot probed with anti-HA antibody and a monoclonal A3H antibody previously designed to recognize human A3H. The monoclonal antibody can recognize some, but not all, of the AGM A3H variants.

One explanation for the difference in antiviral activity is that AGM A3H is not packaged as efficiently into virions as rhesus macaque A3H. To explore this possibility, we measured the amount of A3H in virions and compared the relative abundance between AGM A3H and rhesus A3H. In virions, there is 3.3-fold less AGM A3H (Figure 12D, right), which is reflective of 3-fold less protein observed in cells (Figure 12D, left). Therefore, AGM A3H and rhesus A3H are packaged at similar efficiencies. However, since there is a more than 10-fold difference in viral inhibition, there must be an additional defect on antiviral activity for AGM A3H that is independent of packaging.

In order to more fully explore the relationship between expression levels and antiviral activity of AGM A3H, we codon-optimized both AGM A3H haplotype 1 and rhesus macaque A3H sequences to remove rare codons that might negatively affect protein translation efficiency. Based on codon-usage statistics in primates (see Methods) 100 out of 211 codons were replaced to more frequent codons in the codon-optimized AGM haplotype 1 A3H and 106 out of 211 were replaced in the codon-optimized rhesus macaque A3H. We found that codon-optimization of both AGM A3H and of rhesus macaque A3H increased their expression levels relative to the native codons in each gene (Figure 14A—compare WT to CO (codon-optimized)). Moreover, after codon-optimization, AGM A3H haplotype 1 and rhesus macaque A3H are expressed to similar levels (Figure 14A and Figure 14B, bottom). However, while codon optimization increased the antiviral activity of both AGM and rhesus macaque A3H over that of wild-type (compare Figure 14B to 12C), codon-optimized rhesus macaque A3H still restricts viral infection 10-fold better than codon-optimized AGM A3H against HIVΔvif and 5-fold better against SIVagm.TANΔvif (Figure 14B, top). We also tested virion packaging efficiency of AGM versus rhesus A3H by measuring the relative abundance of each codon-optimized construct in the virion. We found that codon-optimized AGM A3H is packaged with similar efficiency as codon-optimized rhesus macaque A3H (Figure 14C), This shows that while increasing the expression level of the protein is sufficient to improve its antiviral activity, other factors after viral incorporation impact the poor antiviral activity of AGM A3H.

Studies in humans have shown that less active A3H proteins also localize to the nucleus [129, 130]. We therefore asked whether wild-type AGM A3H proteins were expressed to lower levels with low antiviral activity due to a change in localization. Wild-type AGM A3H haplotypes 1, 5, 6, and 10 and rhesus macaque A3H were transfected into HeLa cells and visualized using immunofluorescent microscopy. However, all AGM A3H variants and rhesus macaque A3H were mainly present in the cytoplasm (Figure 14D), demonstrating that a drastic change in localization was not responsible for the lack of potent antiviral activity of AGM A3H.

Our results suggest that the extensive diversity observed in AGM A3H results in lower antiviral activity linked to both protein levels and to other functional differences after packaging due to amino acid divergence between species.

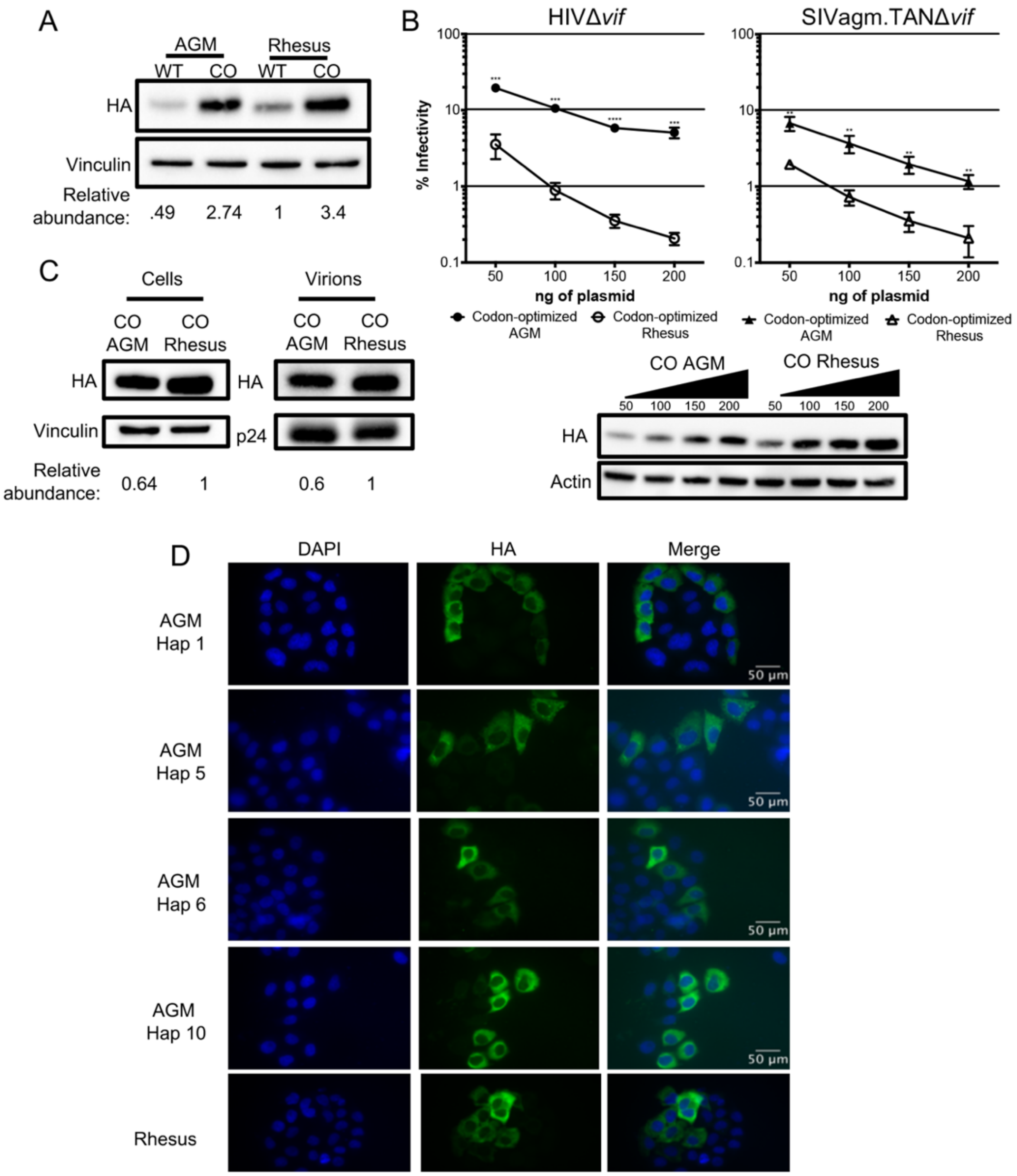

Figure 14. Codon-optimization increases protein expression and antiviral activity
A. Western blot analysis for the expression of AGM A3H haplotype 1, codon-optimized (CO) haplotype 1 A3H, rhesus macaque A3H, and codon-optimized (CO) rhesus macaque A3H. Vinculin was used as a protein loading control. Quantification was done relative to rhesus macaque A3H (normalized to 1). **B.** Top: Single-cycle infectivity assay of HIVΔ*vif* (left, circles) and SIVagm.TANΔ*vif* (right, triangles) in the presence of increasing amounts of A3H plasmid comparing codon-optimized AGM haplotype 1 A3H (black) and codon-optimized rhesus macaque A3H (open). Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. Statistical differences were determined by unpaired t-tests: $^*p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, and $^{****}p \leq 0.0001$. Bottom: Western blot analysis of protein expression level with amounts of plasmid added in panel B. β-Actin is shown as a loading control. **C.** Packaging of A3H proteins into virions analyzed by western blot. Relative abundance in the cellular expression (left) and virion incorporation (right) were determined compared to codon-optimized rhesus macaque A3H. **D.** Subcellular localization of wild-type rhesus macaque and wild-type AGM A3H haplotypes 1, 5, 6, and 10 in HeLa cells. A3H proteins were detected with an anti-HA antibody (green) and DAPI staining was used to detect the nucleus (blue). Images are representative of n = 135 total images over 3 replicates.

Reconstruction of ancestral A3H proteins demonstrates the loss of activity in more recent evolution of AGMs and other primates

We previously explored the evolutionary dynamic of hominoid A3H by reconstructing the ancestor of human/chimpanzee A3H and found that the predicted A3H protein at the human/chimpanzee ancestor had higher antiviral activity than either the extant chimpanzee or human proteins [129]. This suggests that there has been a loss of some activity in both lineages over their evolution. Due to the finding that all tested AGM A3H proteins are poorly active relative to the rhesus macaque A3H, we wanted to reconstruct the ancestral history of A3H leading to the modern AGM lineage. Thus, in order to gain statistical power in the ancestral sequence predictions at each node, we determined the A3H sequence from a broader panel of Old World monkeys in sister clades including De Brazza's monkey, Allen's Swamp monkey, Wolf's guenon, mustached guenon, talapoin, and patas monkey (Figure 15A).

We tested the A3H activity of the closest sister species to the AGMs, patas monkeys, and the A3H activity of a sister species to the rhesus, the sooty mangabey (Figure 15A). Upon transfection into HEK293T cells, the protein expression level of patas monkey A3H was lower in comparison to rhesus macaque A3H, similar to AGM A3H (Figure 15B, bottom). Patas monkey A3H correspondingly had low antiviral activity when tested against HIV∆vif; that is, while AGM and rhesus macaque A3H restricted viral infection 6-fold and 63-fold, respectively, patas monkey A3H restricted HIV∆vif infection only 3-fold (Figure 15B, top). However, active A3H proteins from sooty mangabey and the human A3H haplotype II restrict viral infection 20-fold and 40-fold, respectively. These data show that the antiviral activity is low in a species closely related to AGMs, but a relative of the rhesus macaque and humans encode more active A3H proteins. This finding further suggests that changes in A3H activity may have deeper evolutionary origins in primate evolution.

In order to determine whether A3H antiviral activity was gained in the rhesus macaque/sooty mangabey lineage or was lost in the AGM/patas monkey lineage, we reconstructed the A3H ancestors at various nodes in the Old World monkey phylogeny (Figure 15A). These included the common ancestor of AGMs and patas monkeys (node 1), AGMs, patas monkeys, and its sister clade (node 2), as well as the common ancestor of rhesus macaque, pig-tailed macaque, and sooty mangabey (node 3) and common ancestor of both groups (node 4). Each ancestor was constructed using maximum likelihood with FastML [144] based on the primate species phylogeny. Although the majority of codons were reconstructed with a probability >99%, site 207 was ambiguous in the common ancestor of all tested Old World monkeys and two codons were possible – encoding either an isoleucine (node 4a) or a threonine (node 4b) at position 207. In this case, both possible ancestors were generated using point-mutagenesis.

We then tested the predicted A3H protein at the reconstructed ancestral nodes of the Old World monkey phylogenetic tree for antiviral activity and protein expression level. All A3H ancestors inhibited viral infection between 17- and 33-fold (Figure 15C). Thus, this result suggests that activity was lost within the AGM/patas monkey clade. Similarly, while the A3H node 3 ancestor restricted viral infection 21-fold, rhesus macaque A3H inhibited HIVΔvif by 64-fold, indicating that some activity may have been gained along that lineage. Moreover, the node 1 ancestor representing the common ancestor of AGM and patas monkeys restricted viral infection 33-fold, which was more potent than its descendants, AGMs (6-fold) and patas monkeys (3-fold), despite having similar protein expression levels (Figure 15C). Therefore, loss of activity in A3H in AGM and patas monkeys that occurred after the common ancestor at node 1, which diverged at least 4 million years ago [124], included mutations that both decreased protein expression levels and led to the losses of antiviral activity that are independent of protein expression.

To determine whether differences in packaging of A3H into virions was responsible for the difference in antiviral activity between the node 1 ancestor and an extant AGM A3H, we performed a packaging assay to compare protein amounts in virions. AGM A3H is packaged 2.4-fold less efficiently than the node 1 ancestor (Figure 4D). Moreover, the node 1 ancestor is present at 5-fold higher levels than AGM A3H haplotype 1 in the virion (Figure 15D), which corresponds to the 5-fold better viral inhibition of the ancestor (Figure 15C). This suggests that in the ancestral node directly leading to the AGM lineage, a decrease in packaging of A3H lead to a further loss of antiviral activity.

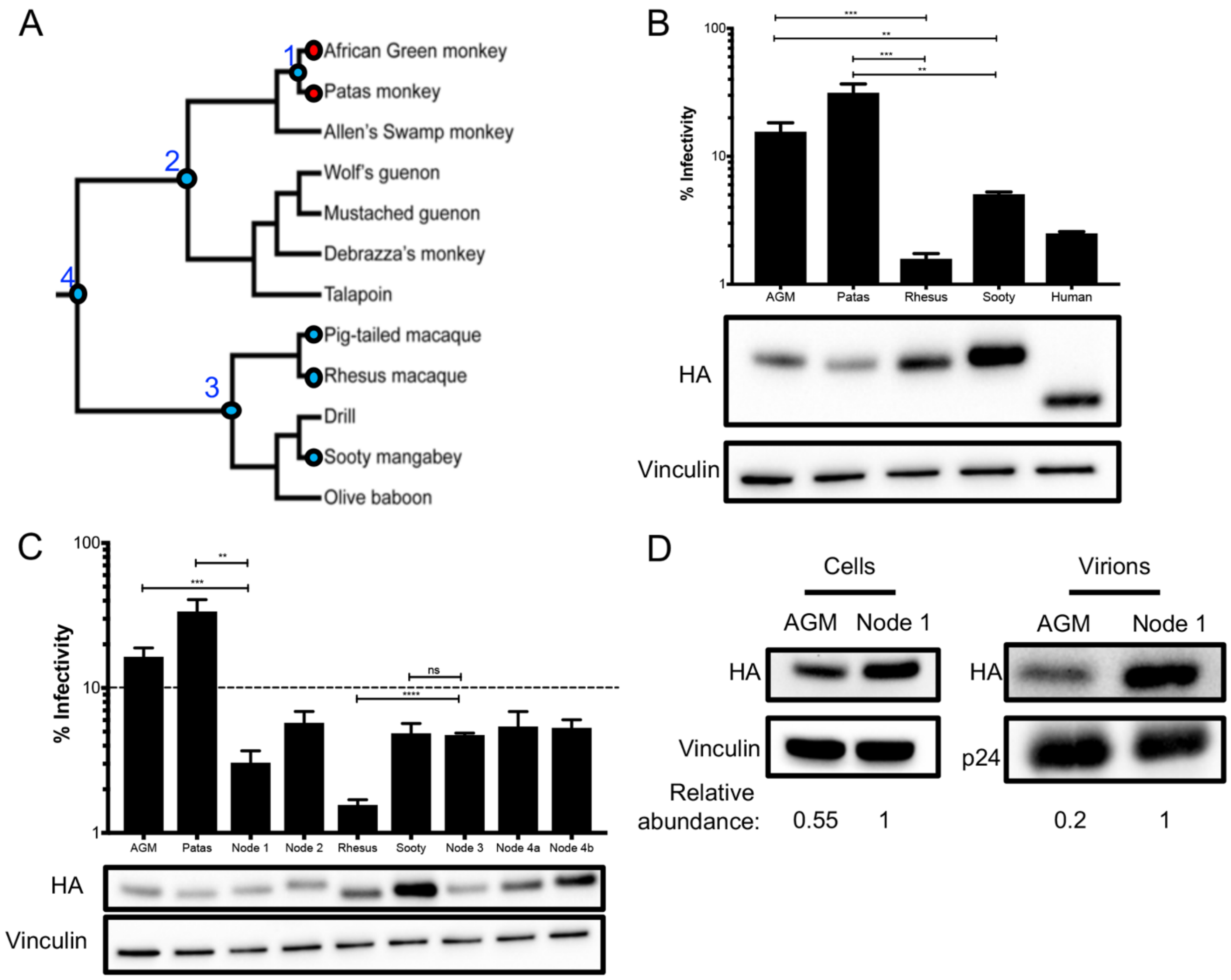

Figure 15. AGM ancestors encode potent antiviral proteins
A. A phylogeny, depicted as a cladogram, based on the accepted species tree of all sequenced Old World primates included in the study [124]. Blue circles denote active antiviral proteins; red circles denote less active antiviral proteins. Ancestral nodes are labeled with numbers (1 – 4). **B.** Top: Single-cycle infectivity assay for HIVΔ*vif* against extant primate A3H proteins. Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. Statistical differences were determined by unpaired t-tests: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, and ****$p \leq 0.0001$. Bottom: Western blot analysis of the protein expression level. Vinculin is used as a loading control. Human A3H is missing 28 C-terminal amino acids due to a natural premature stop codon [127] and thus runs lower on the blot. **C.** Top: Single-cycle infectivity assay for HIVΔ*vif* against ancestral A3H proteins and their extant descendants. Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. Dotted line at 10% is an arbitrary reference point. Bottom: Western blot analysis of the protein expression level. Vinculin is used as a loading control. **D.** Packaging of A3H proteins into virions analyzed by western blot. Relative abundance in the cellular expression (left) and virion incorporation (right) were determined compared to the node 1 A3H ancestor.

Multiple amino acid mutations are responsible for loss of A3H antiviral activity

The ancestral A3H protein at node 1 representing the common ancestor of AGMs and patas monkeys (Figure 15A) has stronger antiviral activity than its extant descendants (Figure 15C). Therefore, we wanted to trace the amino acid mutations that resulted in the subsequent loss along the branches leading from node 1 to AGMs and patas monkeys. The sequence alignment of the node 1 ancestor with AGM haplotype 1 and patas monkey A3H revealed 4 (sites 18, 20, 48, and 171) and 3 (sites 17, 25, and 51) amino acid differences respectively (Figure 16A). One site in AGMs, S20N, created a protein sequence identical to AGM A3H haplotype 8. We introduced the other residues found in the node 1 ancestor into AGM and patas monkey A3H backgrounds to test the expression level and antiviral activity of each mutant.

Notably, no single amino acid mutation in AGM A3H increased the antiviral activity compared to wild-type AGM A3H and all single mutants inhibited viral infection around 5-fold relative to the no A3 control (Figure 16B). In contrast, the double mutants H18R/S20N, H18R/Q48K, and S20N/Q48K, antiviral activity increased to around 13-fold (Figure 16B). Double mutants with the H18R mutation were not statistically significant from the node 1 ancestor, indicating that this amino acid may be most important to rescue antiviral activity. Similarly, patas monkey A3H with single mutations at sites 25 (R25P) and 51 (E51K) did not improve restriction, while a H17R mutation slightly increased the antiviral activity compared to wild-type (3-fold to 5-fold; Figure 16B). However, this single mutation at position 17 did not make patas monkey A3H comparable to the node 1 ancestor. Inserting both H17R and R25P mutations in patas monkey A3H similarly did not further improve restriction (Figure 16B). Taken together, the inability of single point mutations to restore antiviral activity to its ancestral state emphasizes that changes at multiple sites have functional consequences in AGM and patas monkey A3H.

Of interest, site 48 is the only residue that is fixed in all AGMs sequenced for this study. This demonstrates that it occurred first, whereas the other sites are polymorphic and have not yet become fixed in the species (Table 1). To determine whether site 48 alone is sufficient to decrease antiviral activity we added a K48Q mutation in the node 1 background and tested its antiviral activity. However, node 1 K48Q has similar antiviral activity and expression (Figure 16C, bottom), showing that epistasis between multiple amino acids may play an essential role for viral restriction by A3H.

Overall, these data establish that amino acids in the N-terminal portion of A3H are important for antiviral activity. Two out of 3 residues responsible for the loss of activity in patas monkey A3H, sites 17 and 25, are also polymorphic in AGMs (Table 1), suggesting that the loss began at a shared common ancestor not sampled by our analysis. Significantly, many Old World monkeys encode additional mutations near amino acid 15, whose loss in human A3H is known to affect protein stability [129]. For example, residues 18 – 23 have been deleted in De Brazza’s monkey (Figure 16A). Such a large deletion in this region would likely impact the ability of A3H to inhibit viral infection in this species. Similarly, talapoin A3H encodes a histidine at position 18 instead of an arginine, similar to AGM A3H, suggesting that its activity may also be lost (Figure 16A). In contrast, rhesus macaque and sooty mangabey A3H encode active ancestral amino acids at such residues, such as two arginines at positions 17 and 18 (Figure 16A), which may result in the potent antiviral activity of these modern proteins. This implies that loss of A3H function has been lost independently at various points of evolutionary history in both hominoids and Old World monkeys.

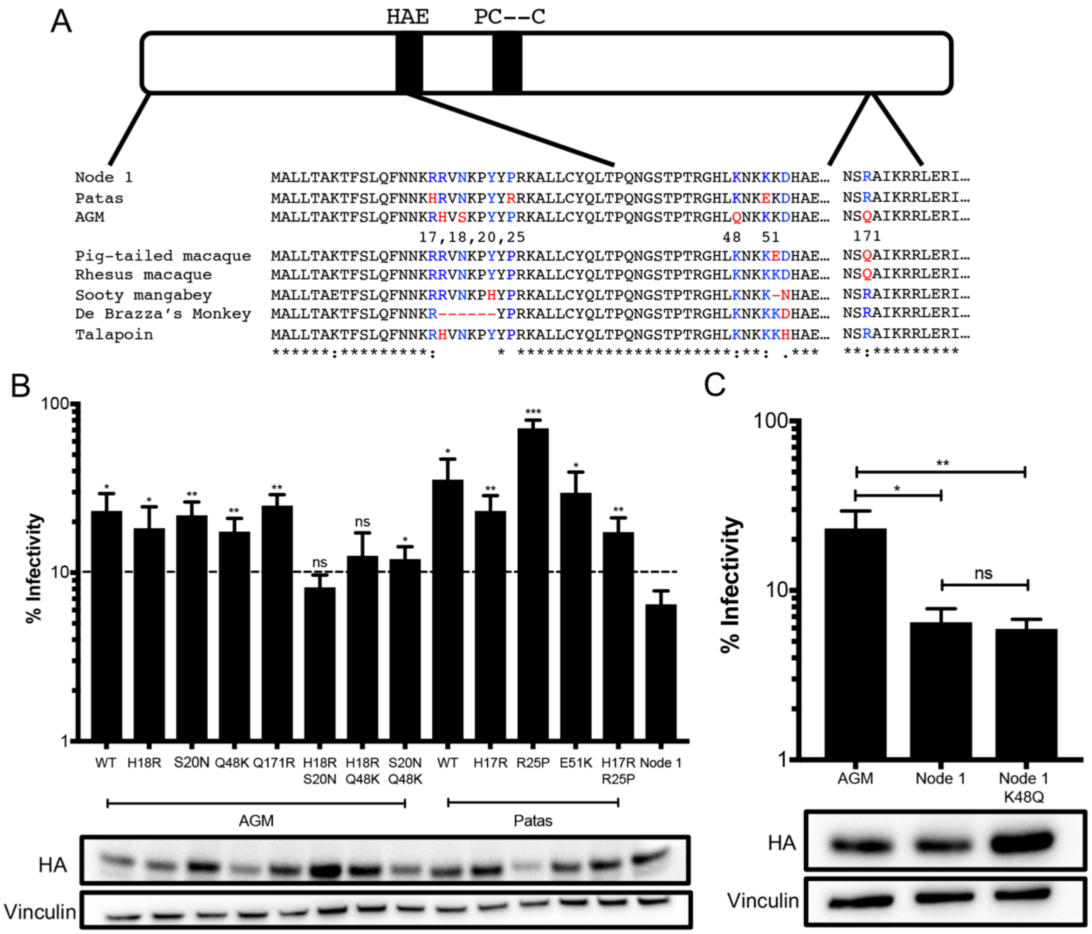

Figure 16. Multiple amino acid mutations required for an increase in antiviral activity
A. Schematic of the A3H protein. Black bars outline the A3H catalytic site. Numbered amino acid residues that are different between AGM A3H, patas A3H, and the node 1 ancestor are outlined in red on the protein sequence alignment, ancestral residues are colored blue. Amino acids that are different in other primates are similarly colored. **B.** Top: Single-cycle infectivity assay for HIV∆*vif* against extant mutants. Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. Statistical differences were determined by unpaired t-tests: $^*p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, and $^{****}p \leq 0.0001$. Statistically significant differences with the node 1 ancestor are depicted as stars above each bar. Bottom: Western blot analysis of protein expression levels of HA-tagged extant mutants made in the AGM and patas A3H backgrounds. Vinculin is used as a loading control. **C.** Top: Single-cycle infectivity assay for HIV∆*vif* against AGM haplotype 1, node 1 ancestor, and node 1 mutant A3H. Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. Statistical differences were determined by unpaired t-tests: $^*p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$, and $^{****}p \leq 0.0001$. Bottom: Western blot analysis of protein expression levels of HA-tagged proteins made in the AGM and patas A3H backgrounds. Vinculin is used as a loading control.

Lack of evidence that evolution in A3H leading to AGM has been driven by Vif

The variability in A3H function throughout Old World monkey evolution suggests an outside selective pressure is driving its loss. Antiviral restriction factors are often rapidly evolving and undergo positive selection [88], which is defined by an excess of nonsynonymous mutations compared to synonymous mutations. Evolutionary conflicts between host restriction factors and viral proteins to either maintain or escape interactions result in an accumulation of nonsynonymous mutations at binding interfaces. A previous study found that A3H is under positive selection [127]. However, the study used fewer primate sequences which can bias the analysis. Therefore, we wanted to re-test positive selection in primate A3H using additional sequences we obtained from Old World monkeys. Using the PAML (phylogenetic analysis by maximum likelihood) program [143], we calculated the number of nonsynonymous mutations (dN) over the number of synonymous mutations (dS) for the entire A3H gene as well as individual codons. In agreement with previous data, we found that the A3H gene is under positive selection and with a dN/dS ratio of 1.3 in all primates. Additionally, models that allow codons to evolve under positive selection fit the data significantly better than models that do not for all primate clades (Figure 17A). We evaluated individual sites within the gene and found that in all primates, a total of 19 residues are under positive selection with a posterior probability > 0.98. One site, position 90, has been implicated in Vif-binding interactions, but only had a posterior probability > 0.98 in one codon model (F3x4; Figure 17A). Since selection can sometimes be driven by the inclusion of specific clades in the analysis, we also condensed our analysis to Old World monkeys alone and found that 5 sites remain under positive selection (blue arrows, Figure 6B), all of which are polymorphic in AGMs (Table 1). No positively selected residues in Old World monkey A3H are in the putative Vif-binding region [154, 155]. Thus, we conclude that multiple sites are under diversifying selection in primates but are likely not driven by lentiviral Vif.

In order to test the hypothesis that the changes in A3H in the lineages leading to the AGMs were not driven by the Vif protein of the lentiviruses that infect AGMs, we tested the restriction capabilities of the node 1 and node 2 ancestors against HIV-1 proviruses expressing Vifs from four SIVagm strains. We also included the codon-optimized AGM A3H since, as it is isolated from an AGM, should be susceptible to degradation by all Vifs encoded by SIVagm. Of interest, the node 2 ancestor encodes an aspartic acid (D) at amino acid 100, which has previously been important for differences in Vif-biding for SIVcpz and HIV-1, while the node 1 ancestor and all AGM A3Hs encode an asparagine (N) [154, 155]. The antiviral activity of all ancestors, as well as codon-optimized AGM A3H, was counteracted by each Vif, resulting in a rescue of viral infection (Figure 17C). Moreover, antagonism of A3H by each Vif protein was accompanied by a loss of packaging of A3H into virions (Figure 17D). Because there are no differences in the ability of Vif to rescue viral infection, these data suggest that Vif is not the primary force on A3H evolution in AGMs, which is instead driven by different selective pressures.

A

# Species	M8a v M8 F3x4 2lnL	p-value	M8a vs M8 F61 2lnL	p-value	M8 BEB P.P. > 0.98
All primates: Hominoids, Old World monkeys, New World Monkeys					
27	69.625	$p < 10^{-17}$	57.926	$p < 10^{-14}$	**7**, **18**, **19**, **20**, **21**, **48**, **53**, 62, 90, **117**, **118**, **127**, **139**, **142**, **152**, **169**, **171**, **172**, **208**
Old World Monkeys					
15	30.051	$p < 10^{-8}$	25.243	$p < 10^{-7}$	**7**, **18**, **53**, **152**, **171**

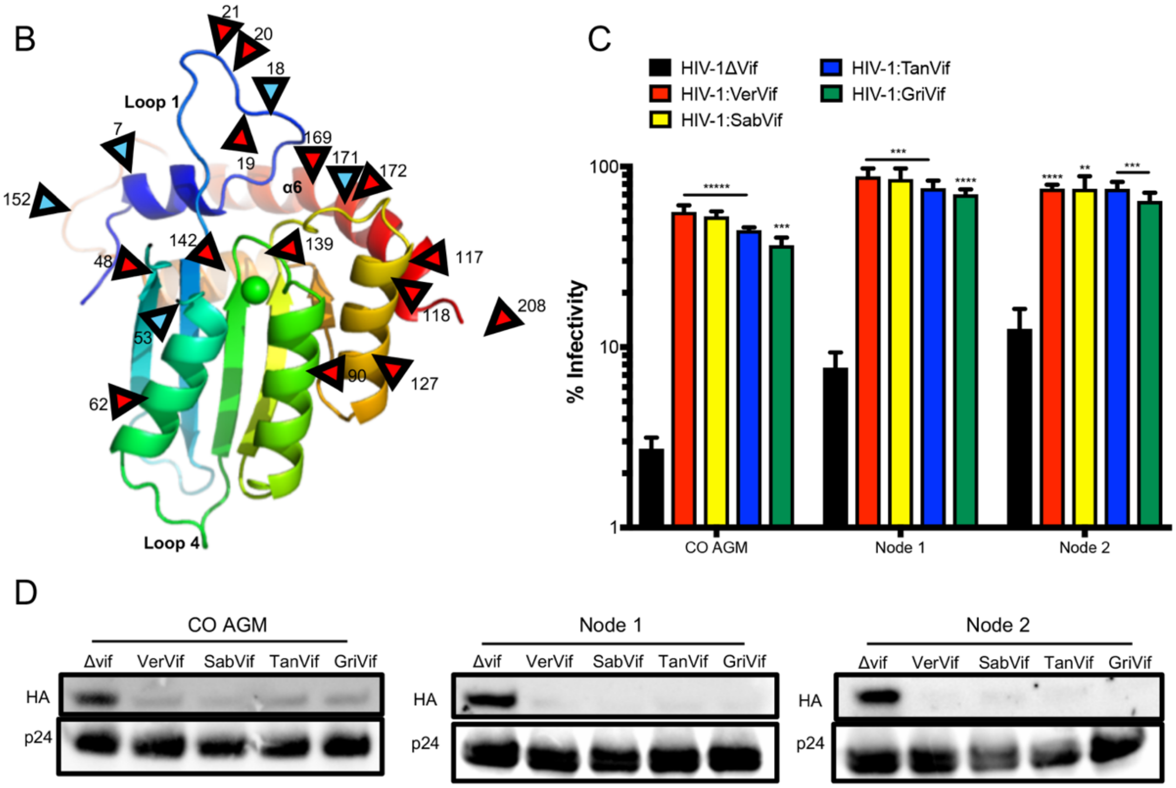

Figure 17. Evolution of A3H is not driven by Vif
A. Results of positive selection analyses of primate A3H. The last column lists sites under positive selection with dN/dS >1 with a posterior probability > 0.98 under M8 Bayes Empirical Bayes (BEB) implemented in PAML model 8. Sites are relative to African green monkey A3H. Sites that had a posterior probability > 0.98 in both codon models (F3x4 and F61) are bolded. **B.** A3H from individual PR01190 was modeled onto the pig-tailed macaque A3H structure (PDB 5W3V). Locations of positively selected sites are denoted by red triangles. Blue triangles denote sites also under positive selection in Old World primates. Site 208 is part of a region missing from the crystallized protein and is therefore not resolved in the model. **C.** Single-cycle infectivity assay done with HIVΔ*vif* and HIV-1 expressing SIVagm Vif in the presence of codon-optimized AGM A3H haplotype 1 (CO AGM), node 1 ancestor, and node 2 ancestor. Relative infection was normalized to viral infectivity in the absence of A3 proteins. Averages of three replicates, each with triplicate infections (± SEM) are shown. Statistical differences were determined by unpaired t-tests: *p ≤ 0.05, **p≤0.01, ***p≤0.001, and ****p≤0.0001. Statistically significant differences to the restriction of HIVΔ*vif* are depicted above each bar. **D.** Western blot analyses demonstrating virion incorporation of A3H proteins in the absence and presence of Vif. p24 is used as a loading control.

Discussion

We show evidence for the recurrent functional loss of APOBEC3H in primates. We found that a decrease in protein expression levels, as well as amino acid mutation in the N-terminal region, results in lower antiviral activity. Using molecular reconstruction of ancestral A3H sequences, we found that the most recent common ancestor of AGMs and patas monkeys likely encoded an active A3H, similar to other common ancestors throughout evolutionary history. This suggests that recurrent loss is a more recent event in primate evolution. Selective pressure by Vif is does not appear to be a primary force behind the evolution of A3H in the AGM clade, but, as loss has occurred both in humans and in other Old World monkeys, there may be a fitness cost to encoding this mutator protein over long evolutionary time periods.

Molecular evolution of A3H protein impacts expression levels and antiviral activity

While increasing the amount of A3H present in cells did increase its capability to inhibit viral infection, greater antiviral activity does not perfectly correlate with higher protein expression. Although codon-optimization increased both the expression level and antiviral activity of A3H, codon-optimized AGM A3H haplotype 1 was still not as potent as codon-optimized rhesus macaque A3H. Thus, both protein expression levels and other amino acid differences that affect the function of A3H have led to the loss of antiviral A3H activity in the AGM lineage. Moreover, although packaging of A3H was equivalent between AGM and rhesus A3H, packaging efficiency of A3H did decrease after AGM A3H diverged from the node 1 ancestor. These results and the results of testing each mutation along the AGM lineage show that multiple mutations were acquired that affected the antiviral functions of the extant AGM A3H protein. Position 48 is fixed in AGMs, indicating that this change likely occurred first, while subsequent amino acid changes concurrently altered the antiviral activity of the protein. However, a K48Q mutation in the node 1 ancestor does not decrease its function (Figure 16C), demonstrating that epistasis between observed amino acid changes more likely lead to its loss. For instance, the AGM A3H double mutant H18R/Q48K inhibits viral infection more comparable to the node 1 ancestor while the individual mutants do not, indicating that both residues are required to gain antiviral activity (Figure 16B). Other AGM A3H haplotypes have also accumulated additional nonsynonymous mutations (Figure 12A), indicating that additional genetic drift may be actively driving A3H to become less antiviral in AGMs, as many animals encode proteins with different expression levels and antiviral activity (Figure 12B and C).

Why has primate A3H maintained partial activity rather than a complete loss?

Loss of protein activity can be indicative of a fitness cost. In the case of human A3H, haplotypes encoding unstable proteins have been linked to greater cancer risk, likely due to its nuclear localization and proximity to host DNA [151]. Moreover, human A3H haplotypes III and IV, which have a deletion at amino acid 15, have little to no antiviral activity. We observe that AGM A3H haplotypes vary greatly in their antiviral restriction and can restrict viral infection anywhere from 17 – 70% (Figure 12A).

However, despite this immense variation in AGM A3H antiviral activity, we observed no large deletions or premature stop codons, which would indicate that the gene itself is being lost. Why has the antiviral activity of A3H not been lost completely in AGMs? It is possible that A3H has been co-opted for a non-lentiviral function in Old World monkeys, thus it is preserved. Alternatively, A3H could be retained through linkage to protective A3G haplotypes. A3G and A3H are located close together on chromosome 19 in AGMs, meaning A3H haplotypes that have ultimately lost most of their function may still be passed onto offspring, particularly if the A3G allele encodes a protein providing a selective advantage. In support of this idea, we noticed poorly active A3H proteins in a monkey encoding protective A3G, as characterized in a separate study [43]. This individual, V005, encodes A3H haplotype 11 as well as an A3G haplotype that cannot be antagonized by Vif proteins from SIVagm.ver or SIVagm.tan. Since A3G is a more potent antiviral and this protein protects individuals from two strains of SIVagm, the protection it supplies may supersede any detrimental effects incurred by the presence of a less antiviral A3H haplotype.

Loss of activity of A3H due to recurrent mutations in a putative RNA binding domain

Two mutations in human A3H gave rise to less potent protein variants: R105G and a deletion of amino acid 15. Of interest, amino acid 15 is positioned within loop 1 of the A3H protein. Recent work has revealed that loop 1, 7, and $\alpha 6$ are important for binding to an RNA duplex [131-133]. We similarly identified that amino acids 18 and 20, found within loop 1, are important for increasing the antiviral activity of AGM A3H (Figure 16B). The analogous locations of such residues may suggest that these changes impact the antiviral activity of A3H by affecting its capability to bind to viral RNA. Dual amino acid changes at residues 18 and 20 in AGMs increase the antiviral activity of A3H close to the levels of a recent common ancestor. An additional mutation close to the catalytic site, Q48K, together with either H18R or S20N in AGMs, further increased antiviral activity. Furthermore, we found that amino acid 17 similarly increased the antiviral activity of patas monkey A3H. This residue is polymorphic in AGMs, suggesting that this change may have occurred in a common ancestor not tested in this study. We also find that other Old World monkeys have changes within loop 1, such as a six amino acid deletion of residues 18 - 23 in De Brazza's monkey and an R18H mutation in talapoin (Figure 16A). The diversity within loop 1 of Old World monkey A3H is indicative that A3H activity was possibly lost multiple times independently. RNA-binding has been shown to play an important role in the antiviral activity of A3H [131, 132] and thus loss of RNA-binding may result in functional loss.

One caveat to our studies is that we have not measured A3H protein levels in primary AGM cells nor in animals. Nonetheless, analysis of activities and protein levels of cloned human A3H genes did mimic their endogenous protein levels and antiviral activity [138].

Vif does not appear to play a role in evolution of A3H along lineage leading to AGMs

AGMs are highly polymorphic in a number of antiviral genes that are specific to infection with lentiviruses in both AGMs explicitly and other primates [125], further highlighting the adaptation to SIV within this particular species. Indeed, we found that multiple sites in A3H are under positive selection in primates, which is suggestive of an evolutionary arms race between A3H and another protein. Evidence of A3H-Vif interactions are evident in hominoids, as SIVcpz Vif adaptation to stable human A3H was crucial for transmission of SIVcpz [139] to humans and HIV-1 Vif is highly variable in its ability to target stable human A3H proteins [134, 136-138]. Interactions between Vif and other A3 proteins have also been well-characterized in chimpanzees and humans [42, 154, 155]. Multiple A3s in chimpanzees protect them from infection by SIVs in Old World monkeys. Therefore, Vif must adapt to antagonize multiple A3 proteins such as A3D and A3G [149]. However, we find that Vif did not play a role in A3H's evolution in AGMs since ancestral proteins were comparably susceptible to antagonism by Vif. Loss of A3H function may facilitate cross-species transmission of SIVagm strains between AGM subspecies [156] due a diminished A3 repertoire. Although many A3 proteins in the family may be redundant, encoding a diverse range of A3s is likely important to achieve maximum protection against lentiviruses.

Since lentiviruses have infected simian primates for millions of years [42], it is unlikely that the lack of residues altering A3H-Vif interactions in AGMs stems from recent infection of Old World monkeys. It is possible that the changes are the result of genetic drift or driven by a different viral pathogen in these primates. Conversely, positively selected residues may indicate evolutionary toggling to preserve or eliminate protein function. This is supported by the finding that many amino acid residues found to increase the antiviral activity of A3H in AGMs are under positive selection (Figure 17A, 6B). It is also possible that the relative importance of different A3 proteins may change dependent on the evolutionary history of a species, driven by the redundancy of the protein family. In the Old World monkeys studied here, less active A3H proteins may impart an increased risk for host genome mutations, thus its function was lost due to the balance between viral protection and host fitness.

Our data implies that A3H function was lost prior to the divergence of different SIVagm strains due to its inactivity in all AGMs. Thus, Vif proteins from the ancestral virus may not have required adaptation to escape the antiviral activity of A3H and instead evolved in response to pressure from the more potent A3G. Expansion of the primate A3 locus provides flexibility of this antiviral protein family to take different trajectories throughout evolution. We have demonstrated that A3H activity is fluid throughout the evolutionary history of primates. In addition to previous work in humans, the A3H homolog in felines, APOBEC3Z3, was recently shown have a similar functional loss [157], demonstrating that loss of A3H and its homolog are frequent throughout a variety of animal species. Such widespread loss of function is suggestive of a potential fitness cost to hosts, although the presence of modern and active A3H proteins exemplifies the importance of encoding a diverse A3 locus in primates.

Chapter 4. Perspectives and Future Directions

In my thesis, I show that A3H is evolutionarily dynamic and activity has been lost multiple times throughout primate evolution. *A3H* is highly polymorphic in four subspecies of African green monkey and all tested protein variants have lost at least some antiviral activity. This decrease in potency is due to both protein expression levels and amino acid mutations that affect protein function. Furthermore, loss of activity in AGMs and closely related patas monkeys was a relatively recent event, as their most recent common ancestor encoded a more active protein. The loss of activity along the AGM lineage was associated with a decrease in packaging efficiency of AGM A3H into virions. I additionally show that the evolution of A3H along the AGM lineage was not driven by Vif, despite strong evidence of positive selection at multiple residues. However, because A3H activity has been lost in some Old World monkey and hominoid species, there may be a fitness cost conferred to the host by more potent A3H proteins in the absence of an outside selective pressure to retain its function. This work contributes to our knowledge of *A3* locus diversity in primates. Having a broad scope of knowledge into how primates control and prevent cross-species transmission can aid in predicting or rapidly treating newly emerging pathogens. Appreciating the fitness landscape of a virus during interactions with its host will also help innovative researchers develop better treatment or cure strategies that can minimize the risk for viral evasion.

A multitude of questions remain concerning the broader implications of APOBEC3 proteins. Analyses regarding mechanisms responsible for the loss of A3H activity, as well as roles of other AGM A3s in viral restriction and the impact of a diverse *A3* locus in the evolutionary arms race with SIVagm Vif, are not yet fully elucidated. Moreover, the field still does not have a full understanding of the importance of A3Z3 proteins, the most evolutionarily divergent in the *A3* family, throughout mammalian history. I will use this discussion to address these topics and explore possible avenues of research.

Mechanisms of APOBEC3H activity loss

A3 proteins are known to interact with RNA, which is an important factor for efficient packaging into virions and antiviral activity. The RNA binding of A3G is the best characterized and has been shown to associate with viral genomic RNAs as well as some cellular RNAs, although there is no consensus as to the extent at which A3 proteins bind cellular nucleic acids. In particular, A3G and A3H have been demonstrated to bind 7SL RNA, a signal recognition particle RNA and important component of the viral ribonucleoprotein complex including Gag and viral genomic RNA [158, 159]. Indeed, the deaminase activity of A3H is RNase-dependent, since the addition of RNase increases deamination of DNA targets, suggesting that DNA and RNA binding to similar surfaces on the protein. Furthermore, a deletion of amino acid 15 (ΔN) in human A3H haplotypes III and IV has been reported to alter the RNA binding activity of the protein, thus reducing its stability and antiviral activity [160].

Recently, the structure of A3H was determined and some groups co-crystalized the protein, from either pig-tailed macaques or humans, in a two-molecule complex bound to an RNA duplex [131, 133]. This RNA duplex is composed of 7 paired bases and 2 unpaired nucleotides at the 5' end. It is not completely certain whether the two proteins contact, but there is evidence showing that residue 23 facilitates an interaction between A3H molecules in humans. Sites involved in RNA binding are clustered in loop 1, loop 7, and the $\alpha6$ helix, in which many positively-charged amino acids form patches. Sites 110 – 115 in loop 7 form an aromatic cage that cap the 3' end of the RNA duplex, while residues 13 - 30 in loop 1 stabilize an RNA kink. A basic patch of amino acids in the $\alpha6$ helix (171, 174, 175, 176, and 179) have also been shown to play a role in RNA binding. Abrogating these patches would conceivably alter the ability of the protein to bind RNA, which may have implications for packaging into virions as well as cellular localization and antiviral activity. For instance, mutants that lack RNA-binding capabilities are found predominantly throughout the entire cell, including the nucleus, while wild-type A3H is mainly seen in the cytoplasm [131]. This supports previous data that unstable human A3H haplotype I localizes more to the nucleus and stable haplotype II is cytoplasmic [130].

Moreover, the connection between RNA and A3 proteins in the context of packaging has not fully been explained. RNA binding is thought to be associated with increased packaging, as interactions between A3G and the nucleocapsid region of HIV-1 Gag are required for efficient incorporation into virions [161-164]. Furthermore, this interaction has been previously shown to be RNA-dependent, although whether this binding is specific [158] or promiscuous [159] remains to be determined. Of note, 7 nt RNA duplex structures similar to that co-crystalized with A3H are enriched in the HIV-1 genome compared to cellular genomic DNA [133]. This seems to indicate that A3H has some form of binding specificity for viral genomes over host genetic material. It is possible that conserved 7 nt motifs, if found across all retroviruses, have driven A3H to preferentially bind these short RNA duplexes, not often observed in cellular DNA, to mediate their specificity and packaging. Alternatively, a different signal upstream of these RNA motifs may facilitate A3H-RNA binding. Evaluating the binding preferences of A3H to other viruses that are predicted to form similar 7 nt RNA duplexes would be useful to determine whether the specificity is truly tied to certain viral RNA structures, other patterns that have not yet been explored, or if it is nonspecific.

While AGM A3H is highly polymorphic, many of the amino acid substitutions observed in the group of analyzed AGMs are located in parts of the predicted structure implicated in RNA binding. Specifically, residue 18 and 20 in loop 1 appear to be important for protein function as creating H18R and S20N mutations together rescue antiviral activity of AGM A3H haplotype 1 to the level of the common ancestor with patas monkeys (node 1; Figure 16B). Characterizing the mechanism for the increase in antiviral activity of this mutant would reveal further evidence that RNA binding is critical for A3H function. If packaging is truly dependent on RNA binding and mutations at residues 18 and 20 abolish this interaction, then an H18R/S20N AGM A3H mutant and the node 1 ancestor should have equal packaging, in contrast to the observed difference between wild-type AGM and node 1 A3H (Figure 15D). However, it is also conceivable that these mutations make A3H a better enzyme, in which case a mutant encoding H18R/S20N would more efficiently edit ssDNA targets. Purifying these proteins and analyzing their activity *in vitro* would determine differences in inherent function.

It is further possible that RNA binding facilitates dimerization of A3H proteins and monomeric forms are quickly degraded in the cell by host pathways, leading to their low expression. Thus, observations that less active haplotypes of human A3H, such as haplotype I, have shorter half-lives may be reflective of an inability to bind RNA rather than another cellular protein that anchors it into the cytoplasm, as some hypotheses predict. Perhaps RNA binding shields sites that would otherwise be tagged for degradation by the proteasome [131]. When the node 1 A3H ancestor and AGM A3H were expressed in packaging experiments, the node 1 A3H ancestor was observed at 2x the level of AGM A3H haplotype 1 in the cell, compared to equal levels seen when A3H alone is overexpressed (Figure 15C vs 15D, left). It is possible that in the absence of viral RNA to bind, node 1 A3H cannot form higher molecular weight complexes and is ultimately degraded by the cell. Assessing the oligomerization and expression of antiviral A3H after viral infection in the presence or absence of RNase compared to less potent A3H proteins would help determine the specific role of RNA in packaging and protein stability.

Overall, further study into why and how A3H proteins display such variation in antiviral function is warranted to better understand its general mechanism of activity. This work would determine how RNA impacts packaging, oligomerization, and enzymatic functions.

What is driving the evolution of A3H?

Although a co-crystal structure of HIV-1 Vif bound to a human A3 protein has not been described, mutagenesis and positive selection studies have extensively characterized their potential binding interfaces. Vif has been shown to drive the evolution of A3G proteins in AGMs and other Old World monkeys [42, 43], indicated by the existence of A3G proteins that are resistant to Vifs from other SIV strains. A3H has been shown to have a similar role in hominoid infections. Stable A3H proteins are a species barrier to SIVcpz infection of humans [139] and has been correlated to lower risk of transmission and slower progression to AIDS [140, 141]. Therefore, *vif* must evolve to overcome stable A3H restriction.

However, the same is not true for AGM A3H. Along the AGM lineage in primate evolution, Vif did not impact the evolutionary trajectory of A3H since all SIVagm Vif variants can rescue viral infection in the presence of A3H proteins dating back to the common ancestor of AGMs and other guenons (node 2). A signal of positive selection is observed in *A3H* using the Old World monkey phylogeny, suggesting that *A3H* is in an evolutionary arms race. One explanation for this is that Vif is driving evolution of A3H in other Old World monkeys, which I did not use in functional tests. An expanded assessment of A3H antiviral activity and susceptibility to Vif across different Old World monkeys may reveal evolution at A3H-Vif binding residues, indicating virus-driven evolution in other primate lineages. Another explanation is that A3H has adopted a different function in the AGM lineage, thus interactions with Vif are irrelevant and its evolution in these monkeys is the result of its new role. For instance, it may play a significant role in viral restriction of a different virus, as A3 proteins have been shown to antagonize DNA viruses such as hepatitis B and herpes viruses [165, 166]. It is possible that AGM A3H is more active against another virus found in AGMs compared to SIVagm, in which case further study into common viral pathogens would be merited.

Alternatively, A3H may be linked to protective haplotypes of other A3 proteins, such as A3G [43]. *A3* genes in AGMs are located in tandem on chromosome 19 and their close proximity makes it likely that genes will be passed on to offspring together. Alleles that provide a selective advantage to such offspring will be more likely persist in populations. Therefore, it can be expected that AGM A3G haplotypes VII or VIII, which protect sabaeus and grivet monkeys from SIVagm.ver/SIVagm.tan infection, may provide a critical advantage to AGMs that encode it, surpassing the negative consequences of a less potent A3H haplotype. Thus, its evolution is driven by beneficial effects to the host from other A3 proteins encoded on chromosome 19 and not necessarily by its interactions with a viral antagonist. A better representation of the *A3* locus composition for individual AGMs by genotyping would demonstrate which alleles are frequently encoded together. Experimental evolution studies to evaluate how *vif* evolves under pressure from multiple encoded *A3s* could also provide useful information regarding not only the importance of having *A3H* but also how multiple A3 proteins impact SIVagm evolution.

What role does A3H play *in vivo*?

Most studies examining the evolution and function of A3H have been conducted *in vitro* or have focused on viruses isolated from single patients. Experimental evolution studies have shown that HIV-1 Vif can evolve to antagonize stable human A3H proteins, suggesting that stable A3H is a relevant restriction factor *in vivo* [137, 138]. Additionally, stable human A3H may have been a barrier to cross-species transmission that required adaptation of SIVcpz *vif* before it could establish infection in humans [139]. However, the importance of A3H has not been fully characterized *in vivo* and it remains unclear how variation in human A3H impacts transmission between individuals and the overall evolution of HIV-1 Vif. Experiments in humanized mice utilizing competition assays and mathematical modeling suggest that there is a robust effect of stable A3H on HIV-1 fitness during replication in the host and that A3H may be a crucial factor in human-to-human transmission [167]. When mice encoding stable haplotypes of human A3H were co-infected with two viral strains, one that encoded *vif* able to degrade stable A3H (hyper) and one with a *vif* that could not (hypo), the virus encoding a hyper *vif* outcompeted hypo *vif* strains. Conversely, in mice encoding human A3H haplotype I, there was no clear advantage conferred to either viral strain. This suggests that the ability to counteract restriction by stable A3H is required for efficient HIV-1 replication, while unstable A3H does not place selective pressures on viral fitness. Furthermore, *de novo* mutations conferring a hyper Vif phenotype, meaning that it can antagonize stable A3H proteins, arose during infection in mice, suggesting that the virus can evolve to counteract A3H restriction *in vivo*. This model could further be used to explore the effects of various A3H haplotypes on viral transmission. Viruses isolated from human patients or those generated in the previous study [167] could be used to infect mice encoding a different haplotype, as it is possible that stable haplotypes of A3H are a transmission barrier to strains isolated from individuals encoding unstable haplotypes.

Questions regarding the evolution of HIV-1 *vif* in response to diverse *A3* repertoires could also be addressed using a human cohort model. Discordant couple cohorts, where one partner is HIV+ and the other is HIV-, have been particularly useful to study treatments preventing the spread of HIV between individuals. Monitoring for seroconversion of the HIV- partner can identify crucial factors for transmission or the effectiveness of new drugs and methods to prevent transmission [168-170]. Such a cohort can also be used to study viral evolution upon infection of a new host. Tracking transmission between individuals encoding different A3H haplotypes would reveal whether HIV-1 *vif* evolves in response to changing *A3H* environments and how often transmission events occur between individuals encoding different *A3H* alleles. Viral strains could be isolated and sequenced from each partner after seroconversion of the HIV- partner to determine the variation in *vif*, where amino acid changes normally associated with A3H antagonism would be identified. Functional assays, such as competition experiments in T cells expressing stable A3H, can further determine advantages or disadvantages to A3H antagonism. It is possible that transmission from an individual encoding less active haplotypes to a partner with an active and stable haplotype is not as successful due to A3H acting as a transmission barrier, while transmission in the opposite direction is more efficient. It is also possible that no relationship between A3H and HIV-1 Vif will be observed, indicating that this virus-host interaction is not actively impacting the evolution of its participants. It would be valuable to assess whether the evolution of HIV-1 *vif* is driven by A3H variants *in vivo* and highlight the role of lentiviral restriction factors in viral transmission and evolution.

Along a similar line of inquiry, one caveat to my thesis is that the work relies on experiments in immortalized cell lines and uses expression vectors rather than measuring endogenous protein. One possibility is that A3H is expressed differently, either at much higher levels or not at all, in the relevant cell types of living AGMs. While studies encompassing characterization of less active human A3H have been extrapolated to human primary cells, it would be worthwhile to measure the effects of endogenous protein on viral replication in AGM primary cell cultures. This would require designing an antibody to recognize a conserved epitope to capture the diversity of AGM A3H proteins and ensure that the protein is indeed expressed in AGM immune cells. I could then compare endogenous protein levels to rhesus macaque primary cells to confirm that they are indeed different. Moreover, primary T cells from infected AGMs may encode proviruses in their genome, which could be sequenced to identify signatures of A3 hypermutation (G-to-A). The ssDNA dinucleotides targeted by A3 proteins are slightly different between A3G and A3D/A3F/A3H. A3G prefers 5’ CC dinucleotides while the others target 5’ TC motifs in minus strand DNA [171]. While A3H cannot be distinguished from hypermutation by A3D or A3F, seeing G-to-A hypermutation from 5’ GA to 5’ AA would hint that other A3 proteins are active in AGM T cells and restricting SIVagm infection. Viral outgrowth assays could also be used to determine whether the proviruses are replication competent or hypermutation by A3D/A3F/A3H is lethal to viral infection. Lastly, genotyping AGMs and following infection could provide associations to different viral outcomes. While AGMs do not progress to AIDS, monitoring acute infection dynamics when exposed to a SIVagm strain from a different subspecies could allow us to determine whether other A3s are linked to protective A3G alleles and view their effects on viral evolution during infection.

Role of other A3 proteins in the AGM A3 – SIVagm Vif evolutionary arms race

AGM A3G has been shown to influence Vif adaptation *in vivo* [43]. After infection of sabaeus monkeys with an SIVagm.ver strain, *vifs* isolated from a homozygous individual encoding protective D130H A3G haplotype VIII successfully rescued some level of infection against haplotype VIII in single-cycle infectivity assays, while Vif proteins from a heterozygous individual (D130/H130) and homozygote (H130) were not. However, Vifs from the individual encoding D130H did not completely rescue viral infectivity to the level of a naturally circulating SIVagm.sab strain. Furthermore, adaptation may have impacted the ability of Vif to recognize other A3G haplotypes in AGMs. This suggests that other A3 proteins may be affecting viral evolution of SIVagm. Since primates encode seven A3 proteins, at least four of which are antiviral, SIVagm Vif must recognize multiple targets during infection. Adaptation of *vif* to maintain antagonistic interactions with all restrictive A3 proteins likely puts constraints on its evolutionary trajectory.

However, since my work has shown that A3H is poorly active in AGMs, other antiviral A3 proteins may play a role, in conjunction with A3G, to drive SIVagm *vif* evolution. While A3G is the most potent antiviral protein in the family and has generally been shown to be the most important in lentiviral restriction, A3F, A3D, or A3C could similarly be highly polymorphic in these animals and be additional selective pressures on viral fitness. For example, chimpanzee A3D is capable of protecting them from cross-species transmissions of SIV strains from Old World monkeys [149]. A survey of other *A3* genes in African green monkeys would reveal an improved representation of the diversity seen within the locus. Functional assays of identified protein variants can determine which A3s are active antivirals and whether any demonstrate species-specific restriction of particular SIVagm strains. If certain A3D haplotypes are found to contribute to species-specificity similar to A3G, for example, it would suggest that A3D and A3G are acting in concert to hinder SIVagm replication. Alternatively, no obvious phenotypes may be observed, indicating that A3G is the primary driver of *vif* evolution in AGMs.

It would also be valuable to expand the current panel of primates, which is currently comprised of 50 animals. Since AGMs are numerous and widely-distributed across Africa, expanding to include more individuals from disparate regions would provide a more complete picture of the diversity encompassed in the *A3* locus. Surveying the genetic diversity of these primates would additionally give a better sense of how such extensive variation influences species-specificity of SIVagm strains, as well as the adaptive forces governing co-evolution of viruses and their hosts.

What can we learn studying AGM populations that are not infected with SIVagm?

Out of six described subspecies of AGM, four are known to be infected with a species-specific strain of SIVagm: vervets, sabaeus, tantalus, and grivets. These primates are infected with SIVagm at a high prevalence: around 70% in adult animals [122]. However, no reports of SIVagm infection have been identified for two other subspecies: malbrouck (*C. cynosuros*) or Bale mountains vervets (*C. djamdjamensis*). This may be indicative of insufficient testing, particularly as Bale mountains vervets are composed of a very small population in the bamboo forests of Ethiopia. Alternatively, these groups could have been infected with a species-specific lentivirus that has since gone extinct, diverged from their sister species before the introduction of SIVagm into the population, or never been exposed to the virus. A recent study evaluated the genetic diversity of AGMs, encompassing the broad distribution and various subspecies [125]. Overall, AGMs were found to be highly diverse and many genes involved in viral processes were under diversifying selection. A number of these genes are specific to SIV infection in AGMs compared to rhesus macaques, suggesting that they may signify host adaptation to SIV. However, because this was a broad survey of the genetic diversity including multiple subspecies of AGM, it would be interesting to further contrast the differences between populations that harbor a characterized SIVagm strain and those that do not. This would require doing more complete SIV infection surveys of malbrouck and Bale mountains vervet groups, as well as taking into consideration new subspecies of AGMs as novel species within the genus are identified. The discovery of new SIVagm strains would be a useful finding for continued SIV research in natural hosts. However, if no new strains of SIVagm are identified, then evaluating whether the genetic diversity seen in vervets, sabaeus, tantalus, and grivet monkeys similarly translates to SIVagm-free populations could provide valuable insight into virus-driven host evolution. An absence of diversity in some antiviral genes would support the idea that lentiviruses have been driving the evolution of AGMs since the introduction of SIVagm. Conversely, observing a similar level of genetic diversity, in the absence of SIVagm infection, would indicate that other pathogens or evolutionary forces have been influencing the antiviral repertoire of African green monkeys. However, this could also signify that these populations have eradicated SIV infection, in which case they would become a useful system to better understand how host adaptation to viral infection can lead to virus extinction. Infections of individuals from SIVagm-free subspecies with current SIVagm strains may reveal disparate immune responses or novel barriers to cross-species transmission. This information could help researchers identify new targets to boost human immune responses to lentiviral infection.

Impact of polymorphic lentiviral restriction factors on SIV infection in other hosts

More than 40 African primates have been found to harbor SIV infection, a handful of which effectively control disease progression, called natural hosts. These natural hosts do not progress to an AIDS-like disease and appear to sustain viral replication without obvious effects on host health. Natural hosts, like AGMs, have provided a plethora of beneficial information that has led to advances in our understanding of HIV biology [52-54]. Thus far, the impact of polymorphisms in *A3G*, *SAMHD1*, and now *A3H* on viral infection have been characterized in AGMs [43, 78]. The protective effects of polymorphic lentiviral restriction factors have also been shown in non-natural hosts like rhesus macaques. SIVmac was the result of a cross-species transmission of SIVsmm from sooty mangabeys. Polymorphisms in A3G likely played an important role in the adaptation of SIVsmm to rhesus macaques [44]. Vif proteins from the SIVmac lineage acquired mutations that specifically allowed antagonism of diverse rhesus macaque A3G proteins, suggesting that divergence in lentiviral restriction factors is important for cross-species transmission and the emergence of new viruses. My work also showed that A3H activity was gained in rhesus macaques (Figure 15C). As macaques are Old World monkeys found in Asia, where natural SIV infections have not been described, this suggests that ancient primate lentiviruses prior to the introduction of SIV drove evolution of the *A3* locus, perhaps after the geographic separation of African and Asian primates. Nevertheless, our understanding of the breadth of diversity within restriction factors is not complete. An examination of restriction factor diversity within African primates harboring lentiviruses would provide a more comprehensive idea of the barriers primate lentiviruses must overcome, even within a single species. The recent publication of the whole genome sequence from a sooty mangabey, which diverged from rhesus macaques in several immune genes, is a step in the right direction [172]. Functional studies should be conducted to determine which genes encode proteins relevant for controlling natural infection. These immune genes should also be sequenced in a broad panel of sooty mangabeys from across Africa to explore whether such genes are diverse in the population or are conserved in the species. Moreover, understanding how this variation may relate to viral evolution would be invaluable to better comprehend how interspecies diversity impacts viral fitness and transmission in natural and non-natural hosts.

Broad importance and evolution of A3Z3 proteins in mammals

The *A3* locus in mammals is highly flexible, with variation in the number and arrangement of different genes. Throughout mammalian history, the locus has expanded via gene duplication and contracted via gene loss and fusion. Of all *A3* genes found in mammals, those with a Z3 domain (A3Z3), like primate *A3H*, are the most evolutionarily divergent [101]. The work presented in my thesis, as well as the work of others, shows that A3H protein function has been lost multiple times in primate evolution: twice in humans, once in AGMs, at least once in patas monkeys, and possibly in other Old World monkeys [129]. Similarly, an A3H homolog in felines lost function in an analogous manner, with a deletion of two amino acids at positions 16 and 17 [157]. However, *A3Z3* genes are found in all currently characterized mammals, including, and not limited to, those from the orders Carnivora (carnivores), Rodentia (rodents), Perissodactyla (odd-toed ungulates), and Chiroptera (bats) [101]. The recurrent loss of A3Z3 function across mammals might suggest that its function is redundant, although no known gene loss has so far been documented over such a long evolutionary time period. Alternatively, A3Z3 proteins may impart a fitness cost on their hosts, as has been documented in some humans encoding less active A3H haplotypes. Incidence of cancer has been linked to unstable A3H haplotype I, which is more nuclear than stable variants and is thus in closer proximity to potential cellular DNA targets [151]. However, my work shows that a loss of A3H function in Old World monkeys is not correlated with localization to the nucleus, suggesting that a different fitness cost may impact AGMs or that the mechanism to mitigate the cost is distinctive. Furthermore, no duplication events or gene fusions involving *A3Z3* genes are known to have taken place, suggesting that either having more than one is detrimental or events are exceedingly rare and have not yet been sampled. Having more complete knowledge of the extent to which mammals encode functional A3Z3 proteins would allow for further study into the various evolutionary trajectories. Functional studies to determine their inherent cytidine deamination as purified proteins, antiviral activity against retroviruses and other viral pathogens, as well as inhibition of retroelements would provide a high-resolution picture of the flux within all *A3Z3* homologs between active proteins and those that may have lost some function. It is possible that some A3Z3 proteins in species not infected with retroviruses may instead carry another inhibitory role, or that they restrict other viruses. Analyzing the functional capabilities of A3Z3 throughout a wide variety of mammals would allow us to explore the consequences, positive or negative, hosts face by encoding these genes.

Concluding remarks

Studying the evolution of virus-host interactions has the power to make fascinating discoveries that are relevant to our basic understanding of viruses, such as viral evolution, as well as translational aspects, like preventing viral transmission. For example, the identification and characterization of innate immune barriers to cross-species transmission could further help us improve predictions concerning which zoonotic pathogens might cause future outbreaks in humans. The more we comprehend about the history of viral infections in a host and the strategies implemented to thwart such an infection, the better we may be able to predict the evolutionary paths that viruses can take. We can then utilize these predictions to design novel therapeutics, such as inhibitors for restriction factor antagonists like lentiviral Vif. Gene therapies have made enormous strides in recent years, particularly with the discovery of CRISPR/Cas9, and are becoming more sophisticated as scientists develop new treatments that can be actively used in patients. Strategies using gene therapy to introduce mutations in antiviral genes that can confer protection against infection could be a successful route towards a functional cure for HIV by engineering cells resistant to infection. Our understanding of how viruses and their hosts interact can help identify potential targets for gene therapy and how viral evolution may impact their effectiveness.

References

1. **Zaitlin M.** 1998. The Discovery of the Causal Agent of the Tobacco Mosaic Disease. World Scientific, Hong Kong.
2. **Stanley WM.** 1935. Isolation of a Crystalline Protein Possessing the Properties of Tobacco-Mosaic Virus. Science **81:**644-645.
3. **Enquist LW, Editors of the Journal of V.** 2009. Virology in the 21st century. J Virol **83:**5296-5308.
4. **Woolhouse M, Scott F, Hudson Z, Howey R, Chase-Topping M.** 2012. Human viruses: discovery and emergence. Philos Trans R Soc Lond B Biol Sci **367:**2864-2871.
5. **De Paepe M, Leclerc M, Tinsley CR, Petit MA.** 2014. Bacteriophages: an underestimated role in human and animal health? Front Cell Infect Microbiol **4:**39.
6. **Chivero ET, Stapleton JT.** 2015. Tropism of human pegivirus (formerly known as GB virus C/hepatitis G virus) and host immunomodulation: insights into a highly successful viral infection. J Gen Virol **96:**1521-1532.
7. **Mills S, Shanahan F, Stanton C, Hill C, Coffey A, Ross RP.** 2013. Movers and shakers: influence of bacteriophages in shaping the mammalian gut microbiota. Gut Microbes **4:**4-16.
8. **Littlejohn M, Locarnini S, Yuen L.** 2016. Origins and Evolution of Hepatitis B Virus and Hepatitis D Virus. Cold Spring Harb Perspect Med **6:**a021360.
9. **Suh A, Brosius J, Schmitz J, Kriegs JO.** 2013. The genome of a Mesozoic paleovirus reveals the evolution of hepatitis B viruses. Nat Commun **4:**1791.
10. **Reed W, Carroll J, Agramonte A, Lazear JW.** 1900. The Etiology of Yellow Fever-A Preliminary Note. Public Health Pap Rep **26:**37-53.
11. **Li W, Wong SK, Li F, Kuhn JH, Huang IC, Choe H, Farzan M.** 2006. Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S-protein interactions. J Virol **80:**4211-4219.
12. **Zhang CY, Wei JF, He SH.** 2006. Adaptive evolution of the spike gene of SARS coronavirus: changes in positively selected sites in different epidemic groups. BMC Microbiol **6:**88.
13. **Wu K, Peng G, Wilken M, Geraghty RJ, Li F.** 2012. Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus. J Biol Chem **287:**8904-8911.
14. **Manz B, Schwemmle M, Brunotte L.** 2013. Adaptation of avian influenza A virus polymerase in mammals to overcome the host species barrier. J Virol **87:**7200-7209.
15. **Rajsbaum R, Albrecht RA, Wang MK, Maharaj NP, Versteeg GA, Nistal-Villan E, Garcia-Sastre A, Gack MU.** 2012. Species-specific inhibition of RIG-I ubiquitination and IFN induction by the influenza A virus NS1 protein. PLoS Pathog **8:**e1003059.
16. **Valmas C, Basler CF.** 2011. Marburg virus VP40 antagonizes interferon signaling in a species-specific manner. J Virol **85:**4309-4317.
17. **Ashour J, Laurent-Rolle M, Shi PY, Garcia-Sastre A.** 2009. NS5 of dengue virus mediates STAT2 binding and degradation. J Virol **83:**5408-5418.
18. **Piot P, Quinn TC.** 2013. Response to the AIDS pandemic--a global health model. N Engl J Med **368:**2210-2218.
19. **Gifford RJ.** 2012. Viral evolution in deep time: lentiviruses and mammals. Trends Genet **28:**89-100.
20. **Deeks SG, Overbaugh J, Phillips A, Buchbinder S.** 2015. HIV infection. Nat Rev Dis Primers **1:**15035.
21. **Barre-Sinoussi F, Chermann J, Rey F, Nugeyre M, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science **220:**868-871.
22. **Gallo R, Salahuddin S, Popovic M, Shearer G, Kaplan M, Haynes B, Palker T, Redfield R, Oleske J, Safai B, et a.** 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science **224:**500-503.

23. **Popovic M, Sarngadharan MG, Read E, Gallo RC.** 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science **224:**497-500.

24. **Van De Perre P, Lepage P, Kestelyn P, Hekker A, Rouvroy D, Bogaerts J, Kayihigi J, Butzler J-P, Clumeck N.** 1984. ACQUIRED IMMUNODEFICIENCY SYNDROME IN RWANDA. The Lancet **324:**62-65.

25. **Piot P, Taelman H, Bila Minlangu K, Mbendi N, Ndangi K, Kalambayi K, Bridts C, Quinn T, Feinsod F, Wobin O, Mazebo P, Stevens W, Mitchell S, McCormick J.** 1984. ACQUIRED IMMUNODEFICIENCY SYNDROME IN A HETEROSEXUAL POPULATION IN ZAIRE. The Lancet **324:**65-69.

26. **Clavel F, Guetard D, Brun-Vezinet F, Chamaret S, Rey MA, Santos-Ferreira MO, Laurent AG, Dauguet C, Katlama C, Rouzioux C, et al.** 1986. Isolation of a new human retrovirus from West African patients with AIDS. Science **233:**343-346.

27. **Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa JD, Arinaminpathy N, Pepin J, Posada D, Peeters M, Pybus OG, Lemey P.** 2014. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. Science **346:**56-61.

28. **Kalish ML, Robbins KE, Pieniazek D, Schaefer A, Nzilambi N, Quinn TC, St Louis ME, Youngpairoj AS, Phillips J, Jaffe HW, Folks TM.** 2004. Recombinant viruses and early global HIV-1 epidemic. Emerg Infect Dis **10:**1227-1234.

29. **Vidal N, Peeters M, Mulanga-Kabeya C, Nzilambi N, Robertson D, Ilunga W, Sema H, Tshimanga K, Bongo B, Delaporte E.** 2000. Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. Journal of Virology **74:**10498-10507.

30. **Letvin NL, Daniel MD, Sehgal PK, Desrosiers RC, Hunt RD, Waldron LM, MacKey JJ, Schmidt DK, Chalifoux LV, King NW.** 1985. Induction of AIDS-like disease in macaque monkeys with T-cell tropic retrovirus STLV-III. Science **230:**71-73.

31. **Chakrabarti L, Guyader M, Alizon M, Daniel MD, Desrosiers RC, Tiollais P, Sonigo P.** 1987. Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature **328:**543-547.

32. **Guyader M, Emerman M, Sonigo P, Clavel F, Montagnier L, Alizon M.** 1987. Genome organization and transactivation of the human immunodeficiency virus type 2. Nature **326:**662-669.

33. **Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain LV, Liegeois F, Loul S, Ngole EM, Bienvenue Y, Delaporte E, Brookfield JF, Sharp PM, Shaw GM, Peeters M, Hahn BH.** 2006. Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. Science **313:**523-526.

34. **Van Heuverswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF, Shaw KS, Takehisa J, Kraus MH, Loul S, Butel C, Liegeois F, Yangda B, Sharp PM, Mpoudi-Ngole E, Delaporte E, Hahn BH, Peeters M.** 2007. Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon. Virology **368:**155-171.

35. **D'Arc M, Ayouba A, Esteban A, Learn GH, Boue V, Liegeois F, Etienne L, Tagg N, Leendertz FH, Boesch C, Madinda NF, Robbins MM, Gray M, Cournil A, Ooms M, Letko M, Simon VA, Sharp PM, Hahn BH, Delaporte E, Mpoudi Ngole E, Peeters M.** 2015. Origin of the HIV-1 group O epidemic in western lowland gorillas. Proc Natl Acad Sci U S A **112:**E1343-1352.

36. **Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemee V, Damond F, Robertson DL, Simon F.** 2009. A new human immunodeficiency virus derived from gorillas. Nat Med **15:**871-872.

37. **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.

38. **Arien KK, Abraha A, Quinones-Mateu ME, Kestens L, Vanham G, Arts EJ.** 2005. The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. J Virol **79:**8979-8990.

39. **Etienne L, Hahn BH, Sharp PM, Matsen FA, Emerman M.** 2013. Gene loss and adaptation to hominids underlie the ancient origin of HIV-1. Cell Host Microbe **14:**85-92.

40. **Bailes E, Gao F, Bibollet-Ruche F, Courgnaud V, Peeters M, Marx PA, Hahn BH, Sharp PM.** 2003. Hybrid origin of SIV in chimpanzees. Science **300:**1713.
41. **Peeters M, Courgnaud V.** 2002. Overview of Primate Lentiviruses and Their Evolution in Non-human Primates in Africa. HIV Sequence Compendium 2002:2 - 23.
42. **Compton AA, Emerman M.** 2013. Convergence and divergence in the evolution of the APOBEC3G-Vif interaction reveal ancient origins of simian immunodeficiency viruses. PLoS Pathog **9:**e1003135.
43. **Compton AA, Hirsch VM, Emerman M.** 2012. The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. Cell Host Microbe **11:**91-98.
44. **Krupp A, McCarthy KR, Ooms M, Letko M, Morgan JS, Simon V, Johnson WE.** 2013. APOBEC3G polymorphism as a selective barrier to cross-species transmission and emergence of pathogenic SIV and AIDS in a primate host. PLoS Pathog **9:**e1003641.
45. **McCarthy KR, Kirmaier A, Autissier P, Johnson WE.** 2015. Evolutionary and Functional Analysis of Old World Primate TRIM5 Reveals the Ancient Emergence of Primate Lentiviruses and Convergent Evolution Targeting a Conserved Capsid Interface. PLoS Pathog **11:**e1005085.
46. **Sawyer SL, Emerman M, Malik HS.** 2004. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol **2:**E275.
47. **Rey-Cuille MA, Berthier JL, Bomsel-Demontoy MC, Chaduc Y, Montagnier L, Hovanessian AG, Chakrabarti LA.** 1998. Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. Journal of Virology **72:**3872-3886.
48. **Diop OM, Gueye A, Dias-Tavares M, Kornfeld C, Faye A, Ave P, Huerre M, Corbet S, Barre-Sinoussi F, Muller-Trutwin MC.** 2000. High levels of viral replication during primary simian immunodeficiency virus SIVagm infection are rapidly and strongly controlled in African green monkeys. J Virol **74:**7538-7547.
49. **Broussard SR, Staprans SI, White R, Whitehead EM, Feinberg MB, Allan JS.** 2001. Simian immunodeficiency virus replicates to high levels in naturally infected African green monkeys without inducing immunologic or neurologic disease. J Virol **75:**2262-2275.
50. **Apetrei C, Gaufin T, Gautam R, Vinton C, Hirsch V, Lewis M, Brenchley J, Pandrea I.** 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.
51. **Mandell DT, Kristoff J, Gaufin T, Gautam R, Ma D, Sandler N, Haret-Richter G, Xu C, Aamer H, Dufour J, Trichel A, Douek DC, Keele BF, Apetrei C, Pandrea I.** 2014. Pathogenic features associated with increased virulence upon Simian immunodeficiency virus cross-species transmission from natural hosts. J Virol **88:**6778-6792.
52. **Klatt NR, Silvestri G, Hirsch V.** 2012. Nonpathogenic simian immunodeficiency virus infections. Cold Spring Harb Perspect Med **2:**a007153.
53. **Pandrea I, Sodora DL, Silvestri G, Apetrei C.** 2008. Into the wild: simian immunodeficiency virus (SIV) infection in natural hosts. Trends Immunol **29:**419-428.
54. **Chahroudi A, Bosinger SE, Vanderford TH, Paiardini M, Silvestri G.** 2012. Natural SIV hosts: showing AIDS the door. Science **335:**1188-1193.
55. **VandeWoude S, Apetrei C.** 2006. Going wild: lessons from naturally occurring T-lymphotropic lentiviruses. Clin Microbiol Rev **19:**728-762.
56. **Bosinger SE, Li Q, Gordon SN, Klatt NR, Duan L, Xu L, Francella N, Sidahmed A, Smith AJ, Cramer EM, Zeng M, Masopust D, Carlis JV, Ran L, Vanderford TH, Paiardini M, Isett RB, Baldwin DA, Else JG, Staprans SI, Silvestri G, Haase AT, Kelvin DJ.** 2009. Global genomic analysis reveals rapid control of a robust innate response in SIV-infected sooty mangabeys. J Clin Invest **119:**3556-3572.
57. **Jacquelin B, Mayau V, Targat B, Liovat AS, Kunkel D, Petitjean G, Dillies MA, Roques P, Butor C, Silvestri G, Giavedoni LD, Lebon P, Barre-Sinoussi F, Benecke A, Muller-Trutwin MC.** 2009. Nonpathogenic SIV infection of African green monkeys induces a strong but rapidly controlled type I IFN response. J Clin Invest **119:**3544-3555.

58. **Gordon SN, Klatt NR, Bosinger SE, Brenchley JM, Milush JM, Engram JC, Dunham RM, Paiardini M, Klucking S, Danesh A, Strobert EA, Apetrei C, Pandrea IV, Kelvin D, Douek DC, Staprans SI, Sodora DL, Silvestri G.** 2007. Severe Depletion of Mucosal CD4+ T Cells in AIDS-Free Simian Immunodeficiency Virus-Infected Sooty Mangabeys. The Journal of Immunology **179:**3026-3034.
59. **Pandrea IV, Gautam R, Ribeiro RM, Brenchley JM, Butler IF, Pattison M, Rasmussen T, Marx PA, Silvestri G, Lackner AA, Perelson AS, Douek DC, Veazey RS, Apetrei C.** 2007. Acute Loss of Intestinal CD4+ T Cells Is Not Predictive of Simian Immunodeficiency Virus Virulence. The Journal of Immunology **179:**3035-3046.
60. **Pandrea I, Silvestri G, Onanga R, Veazey RS, Marx PA, Hirsch V, Apetrei C.** 2006. Simian immunodeficiency viruses replication dynamics in African non-human primate hosts: common patterns and species-specific differences. Journal of Medical Primatology **35:**194-201.
61. **Silvestri G, Sodora DL, Koup RA, Paiardini M, O'Neil SP, McClure HM, Staprans SI, Feinberg MB.** 2003. Nonpathogenic SIV infection of sooty mangabeys is characterized by limited bystander immunopathology despite chronic high-level viremia. Immunity **18:**441-452.
62. **Guindon S, Gascuel O.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52:**696-704.
63. **Sharp PM, Hahn BH.** 2011. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med **1:**a006841.
64. **Bieniasz PD.** 2003. Restriction factors: a defense against retroviral infection. Trends Microbiol **11:**286-291.
65. **Malim MH, Bieniasz PD.** 2012. HIV Restriction Factors and Mechanisms of Evasion. Cold Spring Harb Perspect Med **2:**a006940.
66. **Simon V, Bloch N, Landau NR.** 2015. Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. Nat Immunol **16:**546-553.
67. **Wolf D, Goff SP.** 2008. Host restriction factors blocking retroviral replication. Annu Rev Genet **42:**143-163.
68. **Altfeld M, Gale M, Jr.** 2015. Innate immunity against HIV-1 infection. Nat Immunol **16:**554-562.
69. **Rehwinkel J.** 2014. Mouse knockout models for HIV-1 restriction factors. Cell Mol Life Sci **71:**3749-3766.
70. **Doyle T, Goujon C, Malim MH.** 2015. HIV-1 and interferons: who's interfering with whom? Nat Rev Microbiol **13:**403-413.
71. **Malim MH, Emerman M.** 2008. HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell Host Microbe **3:**388-398.
72. **Nisole S, Stoye JP, Saïb A.** 2005. TRIM family proteins: retroviral restriction and antiviral defence. Nature Reviews Microbiology **3:**799.
73. **Sebastian S, Luban J.** 2005. TRIM5alpha selectively binds a restriction-sensitive retroviral capsid. Retrovirology **2:**40.
74. **Li YL, Chandrasekaran V, Carter SD, Woodward CL, Christensen DE, Dryden KA, Pornillos O, Yeager M, Ganser-Pornillos BK, Jensen GJ, Sundquist WI.** 2016. Primate TRIM5 proteins form hexagonal nets on HIV-1 capsids. Elife **5**.
75. **Sawyer SL, Emerman M, Malik HS.** 2007. Discordant evolution of the adjacent antiretroviral genes TRIM22 and TRIM5 in mammals. PLoS Pathog **3:**e197.
76. **Ballana E, Este JA.** 2015. SAMHD1: at the crossroads of cell proliferation, immune responses, and virus restriction. Trends Microbiol **23:**680-692.
77. **Lim ES, Fregoso OI, McCoy CO, Matsen FA, Malik HS, Emerman M.** 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.
78. **Spragg CJ, Emerman M.** 2013. Antagonism of SAMHD1 is actively maintained in natural infections of simian immunodeficiency virus. Proc Natl Acad Sci U S A **110:**21136-21141.
79. **Fregoso OI, Ahn J, Wang C, Mehrens J, Skowronski J, Emerman M.** 2013. Evolutionary toggling of Vpx/Vpr specificity results in divergent recognition of the restriction factor SAMHD1. PLoS Pathog **9:**e1003496.

80. **Serra-Moreno R, Zimmermann K, Stern LJ, Evans DT.** 2013. Tetherin/BST-2 antagonism by Nef depends on a direct physical interaction between Nef and tetherin, and on clathrin-mediated endocytosis. PLoS Pathog **9:**e1003487.

81. **Lim ES, Malik HS, Emerman M.** 2010. Ancient adaptive evolution of tetherin shaped the functions of Vpu and Nef in human immunodeficiency virus and primate lentiviruses. J Virol **84:**7124-7134.

82. **Hauser H, Lopez LA, Yang SJ, Oldenburg JE, Exline CM, Guatelli JC, Cannon PM.** 2010. HIV-1 Vpu and HIV-2 Env counteract BST-2/tetherin by sequestration in a perinuclear compartment. Retrovirology **7:**51.

83. **McNatt MW, Zang T, Bieniasz PD.** 2013. Vpu binds directly to tetherin and displaces it from nascent virions. PLoS Pathog **9:**e1003299.

84. **Sauter D, Schindler M, Specht A, Landford WN, Munch J, Kim KA, Votteler J, Schubert U, Bibollet-Ruche F, Keele BF, Takehisa J, Ogando Y, Ochsenbauer C, Kappes JC, Ayouba A, Peeters M, Learn GH, Shaw G, Sharp PM, Bieniasz P, Hahn BH, Hatziioannou T, Kirchhoff F.** 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.

85. **Liu Z, Pan Q, Ding S, Qian J, Xu F, Zhou J, Cen S, Guo F, Liang C.** 2013. The interferon-inducible MxB protein inhibits HIV-1 infection. Cell Host Microbe **14:**398-410.

86. **Lu J, Pan Q, Rong L, He W, Liu SL, Liang C.** 2011. The IFITM proteins inhibit HIV-1 infection. J Virol **85:**2126-2137.

87. **Li M, Kao E, Gao X, Sandig H, Limmer K, Pavon-Eternod M, Jones TE, Landry S, Pan T, Weitzman MD, David M.** 2012. Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11. Nature **491:**125-128.

88. **Daugherty MD, Malik HS.** 2012. Rules of engagement: molecular insights from host-virus arms races. Annu Rev Genet **46:**677-700.

89. **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.

90. **Harris RS, Dudley JP.** 2015. APOBECs and virus restriction. Virology **479-480:**131-145.

91. **Stavrou S, Ross SR.** 2015. APOBEC3 Proteins in Viral Immunity. J Immunol **195:**4565-4570.

92. **Cascalho M.** 2004. Advantages and disadvantages of cytidine deamination. J Immunol **172:**6513-6518.

93. **Teng B, Burant CF, Davidson NO.** 1993. Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science **260:**1816-1819.

94. **Greeve J, Altkemper I, Dieterich JH, Greten H, Windler E.** 1993. Apolipoprotein B mRNA editing in 12 different mammalian species: hepatic expression is reflected in low concentrations of apoB-containing plasma lipoproteins. Journal of Lipid Research **34:**1367-1383.

95. **Moris A, Murray S, Cardinaud S.** 2014. AID and APOBECs span the gap between innate and adaptive immunity. Front Microbiol **5:**534.

96. **Gabuzda DH, Lawrence K, Langhoff E, Terwilliger E, Dorfman T, Haseltine WA, Sodroski J.** 1992. Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. J Virol **66:**6489-6495.

97. **Vonschwedler U, Song JP, Aiken C, Trono D.** 1993. Vif Is Crucial for Human-Immunodeficiency-Virus Type-1 Proviral DNA-Synthesis in Infected-Cells. Journal of Virology **67:**4945-4955.

98. **Simon JH, Gaddis NC, Fouchier RA, Malim MH.** 1998. Evidence for a newly discovered cellular anti-HIV-1 phenotype. Nat Med **4:**1397-1400.

99. **Madani N, Kabat D.** 1998. An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. Journal of Virology **72:**10251-10255.

100. **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.

101. **Munk C, Willemsen A, Bravo IG.** 2012. An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. BMC Evol Biol **12:**71.

102. **Hultquist JF, Lengyel JA, Refsland EW, LaRue RS, Lackey L, Brown WL, Harris RS.** 2011. Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. J Virol **85:**11220-11234.

103. **LaRue RS, Andresdottir V, Blanchard Y, Conticello SG, Derse D, Emerman M, Greene WC, Jonsson SR, Landau NR, Lochelt M, Malik HS, Malim MH, Munk C, O'Brien SJ, Pathak VK, Strebel K, Wain-Hobson S, Yu XF, Yuhki N, Harris RS.** 2009. Guidelines for naming nonprimate APOBEC3 genes and proteins. J Virol **83:**494-497.

104. **Aydin H, Taylor MW, Lee JE.** 2014. Structure-guided analysis of the human APOBEC3-HIV restrictome. Structure **22:**668-684.

105. **Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH.** 2003. DNA deamination mediates innate immunity to retroviral infection. Cell **113:**803-809.

106. **Yu Q, Konig R, Pillai S, Chiles K, Kearney M, Palmer S, Richman D, Coffin JM, Landau NR.** 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nat Struct Mol Biol **11:**435-442.

107. **Bishop KN, Holmes RK, Sheehy AM, Davidson NO, Cho SJ, Malim MH.** 2004. Cytidine deamination of retroviral DNA by diverse APOBEC proteins. Curr Biol **14:**1392-1396.

108. **Wissing S, Montano M, Garcia-Perez JL, Moran JV, Greene WC.** 2011. Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. J Biol Chem **286:**36427-36437.

109. **Koning FA, Newman EN, Kim EY, Kunstman KJ, Wolinsky SM, Malim MH.** 2009. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. J Virol **83:**9474-9485.

110. **Mehle A, Strack B, Ancuta P, Zhang C, McPike M, Gabuzda D.** 2004. Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. J Biol Chem **279:**7792-7798.

111. **Jager S, Kim DY, Hultquist JF, Shindo K, LaRue RS, Kwon E, Li M, Anderson BD, Yen L, Stanley D, Mahon C, Kane J, Franks-Skiba K, Cimermancic P, Burlingame A, Sali A, Craik CS, Harris RS, Gross JD, Krogan NJ.** 2011. Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. Nature **481:**371-375.

112. **Desimmie BA, Delviks-Frankenberrry KA, Burdick RC, Qi D, Izumi T, Pathak VK.** 2014. Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. J Mol Biol **426:**1220-1245.

113. **Guo Y, Dong L, Qiu X, Wang Y, Zhang B, Liu H, Yu Y, Zang Y, Yang M, Huang Z.** 2014. Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. Nature **505:**229-233.

114. **Letko M, Booiman T, Kootstra N, Simon V, Ooms M.** 2015. Identification of the HIV-1 Vif and Human APOBEC3G Protein Interface. Cell Rep **13:**1789-1799.

115. **Letko M, Silvestri G, Hahn BH, Bibollet-Ruche F, Gokcumen O, Simon V, Ooms M.** 2013. Vif proteins from diverse primate lentiviral lineages use the same binding site in APOBEC3G. J Virol **87:**11861-11871.

116. **Kouno T, Luengas EM, Shigematsu M, Shandilya SM, Zhang J, Chen L, Hara M, Schiffer CA, Harris RS, Matsuo H.** 2015. Structure of the Vif-binding domain of the antiviral enzyme APOBEC3G. Nat Struct Mol Biol **22:**485-491.

117. **Fomsgaard A, Allan J, Gravell M, London WT, Hirsch VM, Johnson PR.** 1990. Molecular characterization of simian lentiviruses from east African green monkeys. J Med Primatol **19:**295-303.

118. **Hirsch VM, McGann C, Dapolito G, Goldstein S, Ogen-Odoi A, Biryawaho B, Lakwo T, Johnson PR.** 1993. Identification of a new subgroup of SIVagm in tantalus monkeys. Virology **197:**426-430.

119. **Jin MJ, Hui H, Robertson DL, Muller MC, Barre-Sinoussi F, Hirsch VM, Allan JS, Shaw GM, Sharp PM, Hahn BH.** 1994. Mosaic genome structure of simian immunodeficiency virus from west African green monkeys. EMBO J **13:**2935-2947.

120. **Allan JS, Short M, Taylor ME, Su S, Hirsch VM, Johnson PR, Shaw GM, Hahn BH.** 1991. Species-Specific Diversity among Simian Immunodeficiency Viruses from African-Green Monkeys. Journal of Virology **65:**2816-2828.
121. **Fukasawa M, Miura T, Hasegawa A, Morikawa S, Tsujimoto H, Miki K, Kitamura T, Hayami M.** 1988. Sequence of simian immunodeficiency virus from African green monkey, a new member of the HIV/SIV group. Nature **333:**457-461.
122. **Ma D, Jasinska A, Kristoff J, Grobler JP, Turner T, Jung Y, Schmitt C, Raehtz K, Feyertag F, Martinez Sosa N, Wijewardana V, Burke DS, Robertson DL, Tracy R, Pandrea I, Freimer N, Apetrei C, International Vervet Research C.** 2013. SIVagm infection in wild African green monkeys from South Africa: epidemiology, natural history, and evolutionary considerations. PLoS Pathog **9:**e1003011.
123. **Grubb P, Butynski TM, Oates JF, Bearder SK, Disotell TR, Groves CP, Struhsaker TT.** 2003. Assessment of the diversity of African primates. International Journal of Primatology **24:**1301-1357.
124. **Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpler Y, Schneider MP, Silva A, O'Brien SJ, Pecon-Slattery J.** 2011. A molecular phylogeny of living primates. PLoS Genet **7:**e1001342.
125. **Svardal H, Jasinska AJ, Apetrei C, Coppola G, Huang Y, Schmitt CA, Jacquelin B, Ramensky V, Muller-Trutwin M, Antonio M, Weinstock G, Grobler JP, Dewar K, Wilson RK, Turner TR, Warren WC, Freimer NB, Nordborg M.** 2017. Ancient hybridization and strong adaptation to viruses across African vervet monkey populations. Nat Genet **49:**1705-1713.
126. **Duggal NK, Fu W, Akey JM, Emerman M.** 2013. Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations. Virology **443:**329-337.
127. **OhAinle M, Kerns JA, Malik HS, Emerman M.** 2006. Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H. J Virol **80:**3853-3862.
128. **Wang X, Abudu A, Son S, Dang Y, Venta PJ, Zheng YH.** 2011. Analysis of human APOBEC3H haplotypes and anti-human immunodeficiency virus type 1 activity. J Virol **85:**3142-3152.
129. **OhAinle M, Kerns JA, Li MM, Malik HS, Emerman M.** 2008. Antiretroelement activity of APOBEC3H was lost twice in recent human evolution. Cell Host Microbe **4:**249-259.
130. **Li MM, Emerman M.** 2011. Polymorphism in human APOBEC3H affects a phenotype dominant for subcellular localization and antiviral activity. J Virol **85:**8197-8207.
131. **Shaban NM, Shi K, Lauer KV, Carpenter MA, Richards CM, Salamango D, Wang J, Lopresti MW, Banerjee S, Levin-Klein R, Brown WL, Aihara H, Harris RS.** 2018. The Antiviral and Cancer Genomic DNA Deaminase APOBEC3H Is Regulated by an RNA-Mediated Dimerization Mechanism. Mol Cell **69:**75-86 e79.
132. **Ito F, Yang H, Xiao X, Li SX, Wolfe A, Zirkle B, Arutiunian V, Chen XS.** 2018. Understanding the Structure, Multimerization, Subcellular Localization and mC Selectivity of a Genomic Mutator and Anti-HIV Factor APOBEC3H. Sci Rep **8:**3763.
133. **Bohn JA, Thummar K, York A, Raymond A, Brown WC, Bieniasz PD, Hatziioannou T, Smith JL.** 2017. APOBEC3H structure reveals an unusual mechanism of interaction with duplex RNA. Nat Commun **8:**1021.
134. **Li MM, Wu LI, Emerman M.** 2010. The range of human APOBEC3H sensitivity to lentiviral Vif proteins. J Virol **84:**88-95.
135. **Binka M, Ooms M, Steward M, Simon V.** 2012. The activity spectrum of Vif from multiple HIV-1 subtypes against APOBEC3G, APOBEC3F, and APOBEC3H. J Virol **86:**49-59.
136. **Ooms M, Letko M, Binka M, Simon V.** 2013. The resistance of human APOBEC3H to HIV-1 NL4-3 molecular clone is determined by a single amino acid in Vif. PLoS One **8:**e57744.
137. **Ooms M, Brayton B, Letko M, Maio SM, Pilcher CD, Hecht FM, Barbour JD, Simon V.** 2013. HIV-1 Vif adaptation to human APOBEC3H haplotypes. Cell Host Microbe **14:**411-421.
138. **Refsland EW, Hultquist JF, Luengas EM, Ikeda T, Shaban NM, Law EK, Brown WL, Reilly C, Emerman M, Harris RS.** 2014. Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity. PLoS Genet **10:**e1004761.

139. **Zhang Z, Gu Q, de Manuel Montero M, Bravo IG, Marques-Bonet T, Haussinger D, Munk C.** 2017. Stably expressed APOBEC3H forms a barrier for cross-species transmission of simian immunodeficiency virus of chimpanzee to humans. PLoS Pathog **13:**e1006746.
140. **Sakurai D, Iwatani Y, Ohtani H, Naruse TK, Terunuma H, Sugiura W, Kimura A.** 2015. APOBEC3H polymorphisms associated with the susceptibility to HIV-1 infection and AIDS progression in Japanese. Immunogenetics **67:**253-257.
141. **Naruse TK, Sakurai D, Ohtani H, Sharma G, Sharma SK, Vajpayee M, Mehra NK, Kaur G, Kimura A.** 2016. APOBEC3H polymorphisms and susceptibility to HIV-1 infection in an Indian population. J Hum Genet **61:**263-265.
142. **Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD.** 2006. GARD: a genetic algorithm for recombination detection. Bioinformatics **22:**3096-3098.
143. **Yang Z.** 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol **24:**1586-1591.
144. **Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O, Pupko T.** 2012. FastML: a web server for probabilistic reconstruction of ancestral sequences. Nucleic Acids Res **40:**W580-584.
145. **Vermeire J, Naessens E, Vanderstraeten H, Landi A, Iannucci V, Van Nuffel A, Taghon T, Pizzato M, Verhasselt B.** 2012. Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. PLoS One **7:**e50859.
146. **Harris RS, Liddament MT.** 2004. Retroviral restriction by APOBEC proteins. Nat Rev Immunol **4:**868-877.
147. **Belanger K, Savoie M, Rosales Gerpe MC, Couture JF, Langlois MA.** 2013. Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. Nucleic Acids Res **41:**7438-7452.
148. **Pollpeter D, Parsons M, Sobala AE, Coxhead S, Lang RD, Bruns AM, Papaioannou S, McDonnell JM, Apolonia L, Chowdhury JA, Horvath CM, Malim MH.** 2018. Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G. Nat Microbiol **3:**220-233.
149. **Etienne L, Bibollet-Ruche F, Sudmant PH, Wu LI, Hahn BH, Emerman M.** 2015. The Role of the Antiviral APOBEC3 Gene Family in Protecting Chimpanzees against Lentiviruses from Monkeys. PLoS Pathog **11:**e1005149.
150. **Harari A, Ooms M, Mulder LC, Simon V.** 2009. Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H. J Virol **83:**295-303.
151. **Starrett GJ, Luengas EM, McCann JL, Ebrahimi D, Temiz NA, Love RP, Feng Y, Adolph MB, Chelico L, Law EK, Carpenter MA, Harris RS.** 2016. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. Nat Commun **7:**12918.
152. **Burns MB, Temiz NA, Harris RS.** 2013. Evidence for APOBEC3B mutagenesis in multiple human cancers. Nat Genet **45:**977-983.
153. **Wertheim JO, Worobey M.** 2007. A challenge to the ancient origin of SIVagm based on African green monkey mitochondrial genomes. PLoS Pathog **3:**e95.
154. **Nakashima M, Tsuzuki S, Awazu H, Hamano A, Okada A, Ode H, Maejima M, Hachiya A, Yokomaku Y, Watanabe N, Akari H, Iwatani Y.** 2017. Mapping Region of Human Restriction Factor APOBEC3H Critical for Interaction with HIV-1 Vif. J Mol Biol **429:**1262-1276.
155. **Ooms M, Letko M, Simon V.** 2017. The Structural Interface between HIV-1 Vif and Human APOBEC3H. J Virol **91**.
156. **Bell SM, Bedford T.** 2017. Modern-day SIV viral diversity generated by extensive recombination and cross-species transmission. PLoS Pathog **13:**e1006466.
157. **Konno Y, Nagaoka S, Kimura I, Takahashi Ueda M, Kumata R, Ito J, Nakagawa S, Kobayashi T, Koyanagi Y, Sato K.** 2018. A naturally occurring feline APOBEC3 variant that loses anti-lentiviral activity by lacking two amino acid residues. J Gen Virol doi:10.1099/jgv.0.001046.
158. **York A, Kutluay SB, Errando M, Bieniasz PD.** 2016. The RNA Binding Specificity of Human APOBEC3 Proteins Resembles That of HIV-1 Nucleocapsid. PLoS Pathog **12:**e1005833.

159. **Apolonia L, Schulz R, Curk T, Rocha P, Swanson CM, Schaller T, Ule J, Malim MH.** 2015. Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. PLoS Pathog **11:**e1004609.

160. **Zhen A, Du J, Zhou X, Xiong Y, Yu XF.** 2012. Reduced APOBEC3H variant anti-viral activities are associated with altered RNA binding activities. PLoS One **7:**e38771.

161. **Luo K, Liu B, Xiao Z, Yu Y, Yu X, Gorelick R, Yu XF.** 2004. Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging. J Virol **78:**11841-11852.

162. **Schafer A, Bogerd HP, Cullen BR.** 2004. Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. Virology **328:**163-168.

163. **Svarovskaia ES, Xu H, Mbisa JL, Barr R, Gorelick RJ, Ono A, Freed EO, Hu WS, Pathak VK.** 2004. Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. J Biol Chem **279:**35822-35828.

164. **Zennou V, Perez-Caballero D, Gottlinger H, Bieniasz PD.** 2004. APOBEC3G incorporation into human immunodeficiency virus type 1 particles. J Virol **78:**12058-12061.

165. **Suspene R, Aynaud MM, Koch S, Pasdeloup D, Labetoulle M, Gaertner B, Vartanian JP, Meyerhans A, Wain-Hobson S.** 2011. Genetic editing of herpes simplex virus 1 and Epstein-Barr herpesvirus genomes by human APOBEC3 cytidine deaminases in culture and in vivo. J Virol **85:**7594-7602.

166. **Suspene R, Guetard D, Henry M, Sommer P, Wain-Hobson S, Vartanian JP.** 2005. Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. Proc Natl Acad Sci U S A **102:**8321-8326.

167. **Nakano Y, Misawa N, Juarez-Fernandez G, Moriwaki M, Nakaoka S, Funo T, Yamada E, Soper A, Yoshikawa R, Ebrahimi D, Tachiki Y, Iwami S, Harris RS, Koyanagi Y, Sato K.** 2017. HIV-1 competition experiments in humanized mice show that APOBEC3H imposes selective pressure and promotes virus adaptation. PLoS Pathog **13:**e1006348.

168. **Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, Tappero JW, Bukusi EA, Cohen CR, Katabira E, Ronald A, Tumwesigye E, Were E, Fife KH, Kiarie J, Farquhar C, John-Stewart G, Kakia A, Odoyo J, Mucunguzi A, Nakku-Joloba E, Twesigye R, Ngure K, Apaka C, Tamooh H, Gabona F, Mujugira A, Panteleeff D, Thomas KK, Kidoguchi L, Krows M, Revall J, Morrison S, Haugen H, Emmanuel-Ogier M, Ondrejcek L, Coombs RW, Frenkel L, Hendrix C, Bumpus NN, Bangsberg D, Haberer JE, Stevens WS, Lingappa JR, Celum C, Partners Pr EPST.** 2012. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med **367:**399-410.

169. **Celum C, Wald A, Lingappa JR, Magaret AS, Wang RS, Mugo N, Mujugira A, Baeten JM, Mullins JI, Hughes JP, Bukusi EA, Cohen CR, Katabira E, Ronald A, Kiarie J, Farquhar C, Stewart GJ, Makhema J, Essex M, Were E, Fife KH, de Bruyn G, Gray GE, McIntyre JA, Manongi R, Kapiga S, Coetzee D, Allen S, Inambao M, Kayitenkore K, Karita E, Kanweka W, Delany S, Rees H, Vwalika B, Stevens W, Campbell MS, Thomas KK, Coombs RW, Morrow R, Whittington WL, McElrath MJ, Barnes L, Ridzon R, Corey L, Partners in Prevention HSVHIVTST.** 2010. Acyclovir and transmission of HIV-1 from persons infected with HIV-1 and HSV-2. N Engl J Med **362:**427-439.

170. **Lingappa JR, Petrovski S, Kahle E, Fellay J, Shianna K, McElrath MJ, Thomas KK, Baeten JM, Celum C, Wald A, de Bruyn G, Mullins JI, Nakku-Joloba E, Farquhar C, Essex M, Donnell D, Kiarie J, Haynes B, Goldstein D, Partners in Prevention HSVHIVTST.** 2011. Genomewide association study for determinants of HIV-1 acquisition and viral set point in HIV-1 serodiscordant couples with quantified virus exposure. PLoS One **6:**e28632.

171. **Armitage AE, Deforche K, Welch JJ, Van Laethem K, Camacho R, Rambaut A, Iversen AK.** 2014. Possible footprints of APOBEC3F and/or other APOBEC3 deaminases, but not APOBEC3G, on HIV-1 from patients with acute/early and chronic infections. J Virol **88:**12882-12894.

172. **Palesch D, Bosinger SE, Tharp GK, Vanderford TH, Paiardini M, Chahroudi A, Johnson ZP, Kirchhoff F, Hahn BH, Norgren RB, Patel NB, Sodora DL, Dawoud RA, Stewart CB, Seepo SM, Harris RA, Liu Y, Raveendran M, Han Y, English A, Thomas GWC, Hahn MW, Pipes L, Mason CE, Muzny DM, Gibbs RA, Sauter D, Worley K, Rogers J, Silvestri G.** 2018. Sooty mangabey genome sequence provides insight into AIDS resistance in a natural SIV host. Nature **553:**77-81.

Erin I. Garcia

Fred Hutchinson Cancer Research Center
Division of Human Biology and Basic Sciences
1100 Fairview Ave N. Seattle, WA 98109

EDUCATION

University of California – Santa Cruz Santa Cruz, CA
M.S. Science Communication Start date: September 2018

University of Washington Seattle, WA
PhD, Microbiology September 2014 – August 2018
GPA 3.75

Case Western Reserve University Cleveland, OH
NIH Post-Baccalaureate Research Education Program (PREP) Program July 2013 – June 2014
GPA 4.0

New Mexico State University Las Cruces, NM
B.S. Microbiology January 2010 – August 2013
Minor Biochemistry
GPA 3.9, Cum Laude

RESEARCH EXPERIENCE

University of Washington Seattle, WA
PhD Candidate 2015 – Present
Advisor: Michael Emerman
APOBEC3 evolution and function
Thesis title: Ancient loss of APOBEC3H activity in a clade of primates infected with SIV

Case Western Reserve University Cleveland, OH
NIH Post-Baccalaureate Research Education Program (PREP) Scholar 2013 – 2014
Advisor: Eric Arts
Replicative fitness and genetic diversity of HIV-1

New Mexico State University Las Cruces, NM
Howard Hughes Medical Institute Undergraduate Research Scholar 2010 – 2013
Advisor: Kathryn Hanley
Dengue virus evolution in response to mosquito vector immunity
Thesis title: How does mosquito immunity affect dengue virus evolution?

LABORATORY SKILLS

Molecular Biology
    PCR, sanger sequencing, cloning, point mutagenesis, qRT-PCR, Western blot analysis, proteasomal inhibition, pulse-chase analysis.
Virology
    Single-cycle infections in T-cells, virion packaging
Tissue culture
    Transfection, passaging cell culture, immunofluorescence staining
Phylogenetics
    Maximum-likelihood (PhyML) and Bayesian (BEAST) phylogenetic analysis, evolutionary analysis with DataMonkey and PAML

PUBLICATIONS

Garcia, E.I. and M. Emerman. *In press.* Recurrent loss of APOBEC3H activity during primate evolution. J Virol

Garcia, E.I.*, V.R. Montoya*, J. Bennet, C.C. Andrade, K.A. Hanley. The unending host-pathogen arms race: How mosquito RNA interference affects dengue virus evolution. In preparation
*these authors contributed equally

CONFERENCES

ComSciCon-PNW, Seattle, WA	March 16 – 17, 2018
Society for Advancement of Chicanos/Hispanics and Native Americans in Science, Salt Lake City, UT	October 2018
Compelling Science Storytelling: Workshop for Science Communicators, Seattle, WA	April 2017
Conference on Retroviruses and Opportunistic Infections, Seattle, WA	February 2017
Annual Biomedical Research Conference for Minority Students, Nashville, TN	November 2013

PRESENTATIONS

Characterizing APOBEC3H function and evolution in African Green Monkeys — Seattle, WA
Poster presentation, Conference on Retroviruses and Opportunistic Infections — February 2017

Do HIV-1 clone(s) have unique phenotypic properties? — Cleveland, OH
Poster presentation
Biomedical Graduate Student Symposium — May 2014

How does mosquito immunity affect dengue virus evolution? — Las Cruces, NM
Oral Presentation
NMSU Biosymposium — April 2013

TEACHING EXPERIENCE

University of Washington

Seattle, WA
Lecturer

"Research applications" MICROM 431 Prokaryotic Recombinant DNA Techniques	March 2018
"Vector-borne/Zoonotic Diseases" MICROM 301 General Microbiology	May 2018
"Dengue and Zika Viruses" MCB 532 Human Pathogenic Viruses	November 2017
TA/RA Conference Facilitator "Presenting your Research Effectively" and "Being an RA in Health and Biological Sciences"	Fall 2016
Research Mentor, mentored UW undergraduate student Joshua Lee	Summer 2016
Teaching Assistant, MICROM 411 Bacterial Genetics	Winter 2016
Teaching Assistant, MICROM 402 Fundamentals of General Microbiology Laboratory	Fall 2015

New Mexico State University

Las Cruces, NM
Private Tutor, General and organic chemistry, algebra, biology, and human physiology — 2011 – 2014

FELLOWSHIPS, HONORS, AND AWARDS

Young Investigator Scholarship 2017
Travel award to 2017 Conference on Retroviruses and Opportunistic Infections in Seattle, WA

Viral Pathogenesis Training Grant 2015 – 2017
T32AI083203, PI: Overbaugh

NSF Graduate Research Fellowship 2015
Honorable mention

2nd Place Poster Presentation 2014
Biomedical Graduate Student Symposium, Case Western Reserve University

NIH Post-Baccalaureate Research Education Program (PREP) Research Scholar
2013 – 2014
Case Western Reserve University

Best Undergraduate Oral Presentation 2013
Annual Biosymposium, New Mexico State University

Howard Hughes Research Scholar 2010 – 2013
New Mexico State University

Dean's List 2010, 2012, 2013
New Mexico State University

Crimson Scholar 2010 – 2013
New Mexico State University

PROFESSIONAL MEMBERSHIPS

American Society for Microbiology
Northwest Science Writers Association
American Association for the Advancement of Science

SCIENCE WRITING

Blogging at http://www.sciencegonerogue.com/

RELEVANT CLASSES

Writing Courses at the University of Washington
Nonfiction Writing – Craft (Fall 2017); The Art of Writing (Spring 2018)

SCIENCE OUTREACH

Elysian Brewing Great Pumpkin Beer Festival Seattle, WA
Festival sponsored Fred Hutch - communicated science to festival attendees September 2017

Guest Co-Host
Podcast: This Week in Virology - Episode 455 "Pork & Genes" August 2017

Paws on Science Seattle, WA
Pacific Science Center May 2016

Letters to a Pre-Scientist
Letters to a Pre-Scientist Program September 2015 - Present

LEADERSHIP ROLES

Board Member 2017 – Present
Member of Student – Post Doc Advisory Committee (SPAC) at Fred Hutchinson Cancer Research Center

Student Retreat Coordinator 2016 - 2018
Plan and organize UW microbiology student retreat in Friday Harbor, San Juan Islands WA

Recruitment Leader – UW Microbiology Winter 2016
Lead student recruitment effort by organizing activities and student involvement

Student Coordinator for Interdisciplinary Biomedical Hooding Ceremony Spring 2016
Coordinated 1st annual, student initiated, hooding ceremony for PhD students

Student Coordinator for Seminar Series 2015 - 2016
Coordinated casual gatherings with students and weekly seminar speakers after talks in the UW microbiology department