2010

The Impact of Host Factors on Retroviral Evolution and the Identification of a Novel Receptor That Was Used by an Ancient Primate Retrovirus

Steven J. Soll

THE IMPACT OF HOST FACTORS ON RETROVIRAL EVOLUTION AND THE IDENTIFICATION OF A NOVEL RECEPTOR THAT WAS USED BY AN ANCIENT PRIMATE RETROVIRUS

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Steven J. Soll

June 2010

The impact of host factors on retroviral evolution and the identification of a novel receptor that was used by an ancient primate retrovirus

Steven J. Soll, Ph.D.

The Rockefeller University 2010

I

The resurrection of inactive endogenous retroviruses allows us to learn about interactions between extinct pathogens and their hosts that occurred millions of years ago. Two of these paleoviruses, chimpanzee endogenous retrovirus 1 and 2 (CERV1 and CERV2), are relatives of modern murine leukemia viruses that are found in the genomes of a variety of old world primates, but are absent from the human genome. The non-existence of human CERV1 and CERV2 homologues is peculiar given the numerous apparent cross-species transmissions that occurred between ancestors of old world monkeys, gorillas, and chimpanzees. It is possible that antiviral proteins were able to protect human ancestors from colonization by CERV1 and CERV2. Indeed, sequence analyses of modern primate restriction factors has suggested that these genes have evolved under positive selection, presumably due to their combat with invading pathogens throughout primate history. Here we investigate whether TRIM5 and APOBEC3 antiviral factors were able to restrict the replication of CERV1 and CERV2. Such an interaction would imply a potential involvement of these proteins in the limited host range and, perhaps, even the extinction of CERV1 and CERV2. Reciprocally,

activity against CERV1 and CERV2 would suggest that archaic gammaretroviruses contributed to the positive selection observed in TRIM5 and APOBEC3 genes.

Our analyses suggest that TRIM5α proteins did not pose a major barrier to the cross-species transmission or contributed to the extinction of CERV1 and CERV2. However, we uncovered extensive evidence for the inactivation of endogenous gammaretroviruses by the action of APOBEC3 cytidine deaminases. Both CERV1 and CERV2, as well as their homologues in the rhesus macaque, bore mutational scars that are characteristic of APOBEC3 activity. A reconstructed CERV2 Gag was used *in vitro* to confirm that APOBEC3G was capable of restricting CERV2 infection. Therefore, it appears that primate APOBEC3 proteins were capable of targeting ancient primate gammaretroviruses. It remains possible that APOBEC3 proteins were able to limit the cross-species transmission and cause the inactivation of these viruses.

II

Although HIV-1 is able to use an antagonist to defend itself from restriction by APOBEC3 proteins, hypermutation is abundant *in vivo*. It is currently unclear if hypermutants have an impact on the evolution of a viral population. The diversity induced by APOBEC3 may benefit a continually evolving population by providing a large scope of variability for selection to act upon or, quite the opposite, it may constitute a deleterious genetic load. The potential impact that APOBEC3-induced hypermutation can have on a replicating HIV-1 population was studied *in vitro*. After a period of growth in T cell lines, an accumulation of hypermutated sequences was observed in an HIV-1 that lacked the APOBEC3 antagonist, Vif. The resulting viral population

displayed a fitness disadvantage in non-restrictive cells relative to a Vif-proficient virus. In spite of this apparent genetic load, though, the Vif-deficient virus was capable of persistent replication. However, unlike the Vif-positive virus, the hypermutated virus was unable to acquire drug resistance in the presence of nevirapine. Although this result argues against the notion that APOBEC3-induced variability can accelerate HIV-1 evolution, a small degree of APOBEC3-associated mutations also occurred in the Vif-proficient virus and this low-level APOBEC3 activity may have contributed to the emergence of nevirapine resistance.

III

In order to enter cells, retroviruses must usurp a host cell surface receptor. Viral receptors can dictate both the cell tropism and host range of the retroviruses that exploit them. Therefore, the identity of the CERV2 receptor is critical to the study of CERV2 transmissions and endogenizations that took place during ancient primate evolution. To investigate the CERV2 host range and identify its receptor, we constructed a consensus CERV2 envelope protein derived using sequences found in the chimpanzee genome. CERV2 enveloped MLV particles were capable of infecting cell lines from a wide range of species, including humans. The permissivity of human cells for CERV2 argues against the notion that human ancestors were protected from CERV2 by the lack of a functional receptor. Using a Hela cDNA library expressed in CERV2 resistant hamster cells, we identified copper transport protein 1 (CTR1) as a novel retrovirus receptor that was presumably used by CERV2 during its exogenous replication more than one million years ago. Expression of human CTR1 was sufficient to confer CERV2 permissively to

otherwise resistant hamster cell lines, which was accompanied by an increase in virion binding. The observed increase in hamster cell infection that came with CTR1 expression was specific to CERV2 enveloped virus, with no gain in permissivity to pseudotypes bearing envelopes from several other gammaretroviruses. Furthermore, siRNA-induced CTR1 knockdown or $CuCl_2$ treatment specifically decreased CERV2 infection of human cells. We have also identified mutations in highly conserved CTR1 residues that have rendered hamster CTR1 inactive as a CERV2 cell surface receptor, including a deletion in a copper-binding motif that is largely conserved from humans to zebrafish. These receptor-inactivating mutations in hamster CTR1 are accompanied by an increased number of extracellular copper-coordinating residues compared to CTR1 proteins from other species. This apparent compensation may represent an evolutionary barrier that primates would have had to overcome to avoid CTR1 usage by CERV2 or related viruses that might have used this protein as a receptor.

Acknowledgments

I would like to acknowledge my current and past coworkers in the Bieniasz lab for being such a brilliant and inspiring group of people. I have enjoyed working with each of them. I want to especially thank David Perez-Caballero for providing both instruction on many techniques as well as thoughtful support. A large portion of the data presented in Chapter III was the result of a collaborative effort with David. Additionally, I thank Theodora Hatziioannou for her endless advice and the members of the thesis committee for providing invaluable discussions: Nina Papavasiliou, Charles Rice, and Nathaniel Landau.

I also acknowledge the following reagent contributions: Theodora Hatziioannou for human and rhesus macaque APOBEC3G, Nathaniel Landau for chimpanzee APOBEC3G, and Michael Malim for other APOBEC expression plasmids; Stuart Neil for generating the HeLa cDNA library; and David Perez-Caballero for the MLV Gag-GFP expression plasmid.

I am very happy to express my gratitude to my advisor, Paul Bieniasz. His guidance has been integral to all of the work presented in this thesis. I believe that the amount of time that Paul has offered me is unprecedented and I could not imagine a more thorough effort.

Finally, I would like to thank my family and friends because it is thanks to their support that I was able continue through unexpected challenges during the completion of the work discussed here.

Table of Contents

Chapter I. Retroviridae, host proteins, and endogenous retroviruses **1**
Retroviral envelopes and cell-surface receptors *4*
Post-entry events and restriction factor encounters *10*
Hypermutation in vivo *17*
Post-reverse transcription events and production of viral progeny *21*
Endogenous retroviruses and the lessons that they can offer *24*

Chapter II. Materials and Methods **29**
Virions and cell lines *29*
Full-length HIV-1 *32*
Endogenous capsid sequence retrieval *33*
Generation of endogenous retrovirus CA-NTD/MLV chimera libraries *35*
Sequence analysis *37*
Gene Synthesis *38*
APOBEC restriction assays *40*
HIV evolution in vitro *41*
Receptor Screen *43*
Virion Binding *44*
RNA-interference *45*

Chapter III. Were TRIM5 and APOBEC3 proteins capable of restricting CERV1 and CERV2? **47**
CERV1 and CERV2 capsid N-terminal domains *47*
MLV Chimeric Viruses containing primate and mouse endogenous gammaretroviral CA-NTDs are generally resistant to TRIM5α proteins *49*
Evidence of endogenous gammaretrovirus restriction by APOBEC3 family members *54*
Confirmation of mutational biases in endogenous gammaretroviruses *59*
Biased patterns of G to A mutation in primate ERVs are not explained by spontaneous cytidine deamination *65*
Effects of cytidine deamination on ancient primate gammaretrovirus CA-NTD function *74*
CERV2 restriction by APOBEC3 proteins in vitro *76*

Chapter IV. The impact of APOBEC3 proteins on HIV-1 evolution in vitro **84**
Replication of HIV-1 delta Vif in permissive and semi-permissive cell lines *84*
Selection of drug resistance in Vif+ virus *91*
Selection of an escape mutant using patient plasma with neutralizing activity *100*

Chapter V. Identification of the CERV2 Cellular Receptor **108**
CERV2 entry into primate cells requires CTR1 and occurs at the plasma membrane *118*
CTR1 induces binding of CERV2 to the plasma membrane *123*
Genetic determinants of CTR1 receptor function *124*

Chapter VI. Discussion **130**

References **143**

List of Figures

Figure 1. Phylogeny of retroviruses 2
Figure 2. The retrovirus replication cycle 5
Figure 3. Reverse Transcription 13
Figure 4. Retroviral integration 22
Figure 5. Phylogeny of chimpanzee endogenous retroviruses (CERVs) 27
Figure 6. Phylogenetic tree of ancient primate gammaretroviral capsid NTDs 48
Figure 7. Resistance of chimeric MLVs containing primate gammaretroviral CA-NTDs to primate TRIM5 proteins 53
Figure 8. Analysis of ancient ERV CA-NTDs 57
Figure 9. Phylogenetic tree illustrating the relationship between endogenous primate gammaretroviral Env sequences and other gammaretroviral Envs 60
Figure 10. Analysis of G to A changes in CA-NTD and Env sequences in primate gammaretroviruses 63
Figure 11. Examples of the RhERV2 proviruses illustrating variation in type, burden, and distribution of G to A mutations among endogenous primate gammaretroviruses 65
Figure 12. Comparative analysis of the context in which G to A versus C to T changes occur in primate gammaretroviral CA-NTDs 69
Figure 13. Comparative analysis of the burden of, and the context in which, G to A versus C to T changes occur in endogenous primate and murine gammaretroviral Env sequences 72
Figure 14. Comparative analysis of the burden of, and the context in which, G to A versus C to T changes occured in CERV2 proviruses 73
Figure 15. Effects of cytidine deamination on CA-NTD function 76
Figure 16. A functional CERV2 Gag consensus and its sensitivity to human APOBEC3G 80
Figure 17. Restriction of CERV2 Gag-containing virions by human, chimpanzee, or rhesus macaque APOBEC3G 82
Figure 18. Lack of potent restriction of CERV2Gag-MLVPol by human APOBEC proteins other than APOBEC3G 83
Figure 19. Replication of HIV with or without Vif in MT2 or CEMSS cells 86
Figure 20. Detection of APOBEC3-induced mutations 90
Figure 21. Inhibition of spreading infection in MT2 cells by antiretroviral drugs 92
Figure 22. Single and dual-drug selection in MT2 cells 95
Figure 23. Evolution of nevirapine but not AZT resistance 96
Figure 24. Nevirapine resistance mutations 98
Figure 25. Selection with Nevirapine in CEMSS cells 99
Figure 26. Selection using patient blood plasma with virus-neutralizing activity 102
Figure 27. Escape mutants show decreased sensitivity to neutralization and remain X4-tropic 103
Figure 28. A V3 loop mutation was fixed into viral population during selection by neutralizing plasma 104
Figure 29. Clustal alignment of envelope protein sequences 110

Figure 30. Human CTR-1 is sufficient to confer permissivity to CERV2 env mediated entry in hamster cells 117
Figure 31. CTR1 is necessary for CERV2 infection in human cells 119
Figure 32. CERV2 entry is pH-independent 120
Figure 33. Human CTR1 promotes binding of CERV2 particles to CHO cells 122
Figure 34. Genetic determinants of CTR1 viral receptor function 127
Figure 35. Clustal alignment of CTR1 proteins 129

List of Tables

Table 1. Drug resistance mutations that could result from APOBEC3 activity 91
Table 2. Number of neomycin-resistant colonies from CHO-PGSA cell expression libraries 114
Table 3. Surface staining of CHO cells expressing N-terminal HA tagged CTR1 128

Chapter I. Retroviridae, host proteins, and endogenous retroviruses

The two processes that define retroviruses are the synthesis of a DNA intermediate from two single stranded RNA genomes, and the subsequent integration of the viral genome into host chromosomal DNA (1). The presence of provirus in cellular genomes of the host means that the viral genome will be copied by cellular DNA replication machinery and then carried forward into progeny cells. The long-term residence of human immunodeficiency virus (HIV) in target cells, including long-lived resting CD4+ T cells, is likely to be a primary reason why patients are unable to clear the virus even during highly active antiretroviral therapy (2-4). Retroviral integration into the genomes of target cells also allows for the possibility of Mendelian inheritance of proviruses if their integration occurs in the host germ line (5). This route of vertical transmission results in an organism-wide establishment of provirus in the genome of progeny. Such endogenization events have occurred numerous times during primate evolution and, in some cases, are shared by diverse primate lineages (6-12). Endogenous retroviruses thus represent a record of ancient retroviral infections and these paleoviruses can provide us with information regarding past interactions of hosts and pathogens (13).

Retroviruses are categorized into several genera based on their phylogenetic relatedness (Figure 1). The lentiviruses include HIV and the simian immunodeficiency viruses (SIV). At least three transmissions of an SIV from chimpanzees (SIVcpz) resulted in HIV-1; a pathogen whose infection causes acquired immunodeficiency

syndrome (AIDS) and, according to the World Health Organization, has spread to more than 60 million people worldwide (14, 15). Lentiviruses are complex retroviruses in that they encode several small accessory genes in addition to the GagPol and Env genes common to all retroviruses. In contrast, the lack of accessory genes in gammaretroviruses classifies them as simple retroviruses. The gammaretrovirus genus includes the murine leukemia viruses (MLV), feline leukemia virus (FeLV), and gibbon ape leukemia virus (GALV).

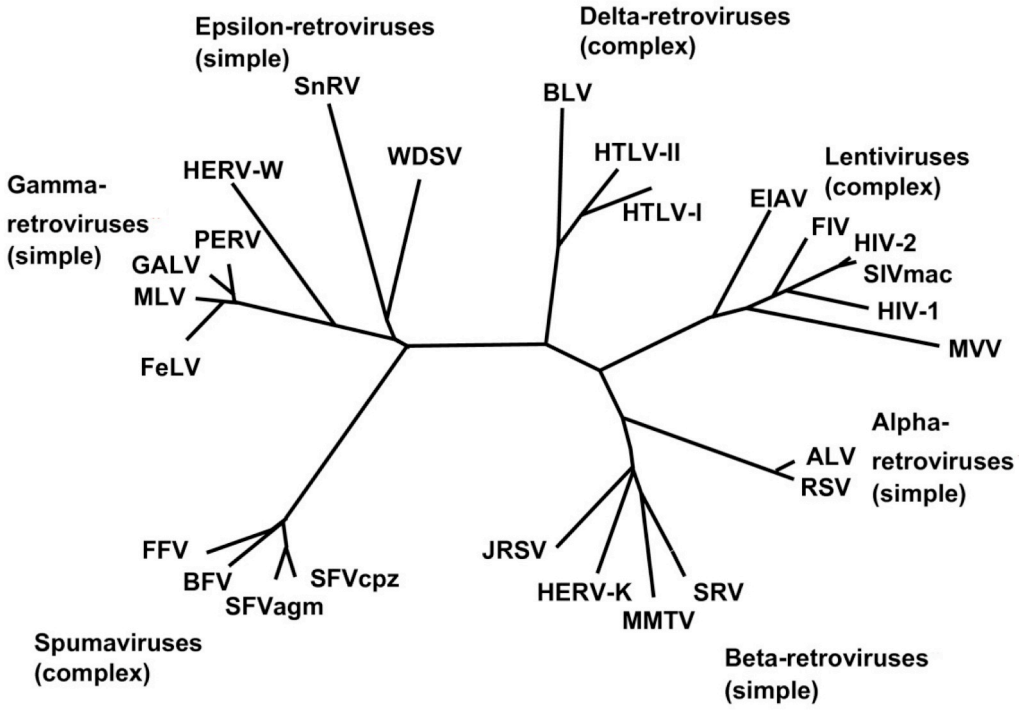


Figure 1. Phylogeny of retroviruses. An Unrooted neighbor-joining tree based on an alignment of reverse transcriptase amino acid sequences is shown. Adapted from Weiss *Retrovirology* 2006 (16).

Lentiviruses have rarely been observed to be endogenous in their hosts. The only known endogenous lentiviruses were identified in the genomes of the European rabbit and the gray mouse lemur of Madagascar (17, 18). A phylogeny of reverse transcriptase

sequences and the presence of the Tat and Rev accessory genes suggest that the rabbit endogenous lentivirus (RELIK) constitutes a unique lentivirus subgroup. The existence of orthologous RELIK insertions in the European hare suggests that these viruses replicated before the divergence of the European hare and rabbit from a common ancestor 12 million years ago (19). Phylogenetic analyses using conserved portions of lentiviral GagPol grouped the gray mouse lemur endogenous lentivirus with the modern primate lentiviruses and it was suggested that this virus represents an evolutionary intermediate between the SIVs and ancestral lentiviruses. Importantly, the identification of an SIV in the lemur suggests that either primate lentiviruses were able to spread in a vector species capable of traversing water, or primate lentiviruses are as old as the last colonization of Madagascar by terrestrial mammals 14 million years ago.

Unlike lentiviruses, endogenous gammaretroviruses are common in diverse mammalian species and these viruses include the porcine endogenous retroviruses (PERV), feline endogenous retroviruses (RD114 and endogenous FeLV), and a variety of endogenous MLVs (all discussed in later sections). Gammaretroviruses are also abundant in primate genomes and, among them, chimpanzee endogenous retroviruses 1 and 2 (CERV1 and CERV2) are the groups most closely related to the modern murine leukemia virus (8, 9, 20). CERV1 and 2 are of particular interest because, in spite of the existence of homologues in the genomes of bonobos, gorillas, and old world monkeys, they are curiously absent in humans. Thus, it appears that both viruses replicated following the divergence of the human and chimpanzee lineage six million years ago, with cross species transmissions resulting in endogenizations within diverse old world primate species, but not in the ancestors of modern humans. It remains possible that a

genetically encoded protection guarded human ancestors from CERV1 and CERV2 invasion during the time that they were becoming endogenized in the non-human primate lineages 1.3 to 6.0 million years ago (8). An exploration of such possibilities, though, warrants an initial discussion of the ways in which retroviral replication depends on some host factors and is assaulted by others.

Retroviral envelopes and cell-surface receptors

Infection of a cell with a retrovirus particle begins with the interaction of the viral envelope glycoprotein with a cell surface receptor (Figure 2). The maturation of the envelope involves furin cleavage during its transport through the Golgi to form a surface protein (SU) and a single-pass transmembrane protein (TM) (21, 22). SU contains receptor-binding regions, while TM is responsible for fusion of viral and cellular membranes. In HIV, trimeric gp120 binds to receptor molecules and sits atop trimeric gp41 (23), both of which are known to be the target of antibodies that are capable of neutralizing the virus [reviewed in (24, 25)]. Although monoclonal antibodies have been isolated whose recognition of antigens on the envelope surface is dependant on glycosylated residues (26, 27), glycans are likely to generally increase the difficulty in acquiring antibody recognition of key envelope structures (28). Indeed, mutagenesis at glycosylated residues generally increases the neutralization sensitivity of HIV-1 envelope to polyclonal sera (29-31). In addition, evolution by the virus repeatedly allows it to persist in the face of continual antibody generation by the host (32-34). gp120 contains five variable loops and sequence flexibility in these regions, along with alterations in

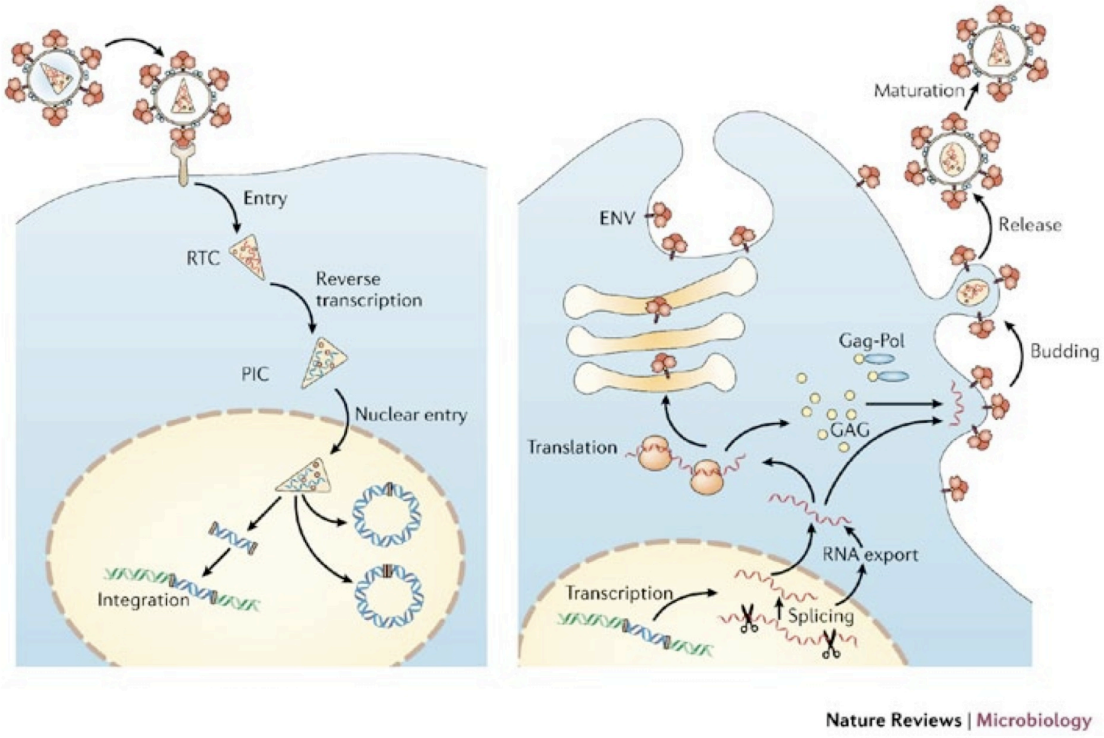


Figure 2. The retrovirus replication cycle. The left panel depicts steps that take place during retroviral infection and the right panel describes the production of virus particles. The various stages of the cycle are referred to throughout this chapter. Taken from Goff *Nature Reviews Microbiology* 2007 (35).

gp120 glycosylation, aid the virus in its escape from immune responses (23, 34, 36-39).

HIV-1 cell tropism is greatly limited by the use of CD4 as its receptor, which is primarily expressed only on helper T-cells, macrophages, and dendritic cells (40). Target cell range is further narrowed by the requirement for gp120 to also bind to either CCR5 or CXCR4 chemokine receptors (41-44). The identification of two HIV-1 co-receptor usages largely explained the ability of only some isolates to induce cell fusions (or syncytia) in laboratory T cell lines, in that the ability to induce T cell syncytia was associated with CXCR4 usage (45-47). Intriguingly, most transmitted viruses use CCR5 (R5-tropic) and about 50% eventually undergo a shift to CXCR4 usage (X4-tropic) that is

temporally correlated with disease progression (48, 49). Determinants of this differential co-receptor usage are largely found in the V3 loop of gp120 (50-55). It appears that the co-receptor utilized by HIV-1 can dictate which $CD4^+$ T cell compartment it predominantly infects. Studies in rhesus macaques using a chimeric SIVmac containing either X4-tropic or R5-tropic HIV-1 envelopes (X4 or R5 SHIVs) have demonstrated that X4-tropism causes the preferential depletion of naïve CD4+ T cells in the peripheral blood and in solid lymph tissue, while R5 SHIVs primarily deplete memory T-cells in the gut (56-58). Further analyses confirmed that biased patterns of infection correlated well with the depletion of specific cellular compartments.

It remains unclear, though, whether a shift in target cell range during the emergence of X4-tropism from initial R5-tropism causes disease progression or if, quite the opposite, the depletion of target cell availability during disease progression forces tropism change. In either case, there likely exists a selective pressure that prevents or delays the emergence of X4-tropic variants. Indeed, change from an R5- to an X4-tropism during SHIV infection in rhesus macaques have a propensity to occur in animals with relatively weak immune responses and the resulting X4-tropic viruses typically display an increased susceptibility to neutralization by both polyclonal and monoclonal antibodies (59-62). These observations support the notion that immune system abnormalities are just as likely to be a cause rather than an effect of X4-tropic virus emergence.

In addition to defining a viral pathogen's specificity for certain cell types within an infected host, receptor usage is also crucial to the ability of a virus to transmit between species. Indeed, cellular receptors are often a primary determinant of gammaretroviral

host range. In the case of MLV, five groups of viruses exist that each display a particular receptor usage and host range (termed interference groups) (63). It is thought that two variable regions in MLV SU that surround a conserved core of β-sheets are responsible for determining receptor specificity (64-66). Ecotropic MLVs use the mouse cationic transport protein (CAT-1) for entry and display a strictly murine host range (67). In contrast, an amphotropic envelope can allow MLV to replicate in cells from a variety of mammalian species through a shift in their receptor usage to the inorganic phosphate transporter, PIT-2 (68-70). A variant of amphotropic MLV, 10A1, is able to use both PIT-2 and the closely related PIT-1, placing it in an interference group that includes gibbon ape leukemia virus (GALV) and feline leukemia virus (FeLV) group B (71).

Similar to the host range expansion allowed by amphotropic envelope, MLV recombinants carrying envelopes from xenotropic or polytropic endogenous MLVs are able to use the xenotropic/polytropic receptor (XPR-1), allowing for distinct host ranges that are determined by their differing interactions with the same receptor (72-77). Chimeras of human and hamster XPR1 demonstrated that the third extracellular loop from human XPR1 allows binding of both xenotropic and polytropic MLV, while the inclusion the second human extracellular loop results in a receptor that is specific for xenotropic MLV (78). Although polytropic MLV generally has a broader rodent tropism compared to its xenotropic relative, some wild mice do express XPR1 proteins capable of allowing xenotropic MLV infection (79, 80). It is tempting to speculate that the observed mouse XPR1 variability is the product of selective pressure imposed by MLV.

The determination of species-specific tropism by envelope and receptor interaction is also a common theme among gammaretroviruses other than MLV. For

example, it has been posited that mutations in the GALV envelope allowed for an expansion of zoonotic potential that led to koala retrovirus, an endogenous retrovirus that continues to replicate exogenously in wild koalas (81, 82). In the case of FeLV and PERV, both are divided into three groups, each with a distinct receptor usage that determines particular species tropisms. Interestingly, all FeLV and PERV receptors that have been identified are transport proteins. FeLV group A is transmitted between felines using the cell surface receptor, thiamine transporter 1 (THTR-1) (83). Within the new host, FeLV-B and/or FeLV-C pathogenic variants emerge that, unlike their FeLV-A predecessor are able to infect canine and human cells. FeLV-B arises through recombination with endogenous retroviruses and consequently has a receptor usage that is shifted to feline PIT-1 or PIT-2 and is also capable of using human PIT-1 (84-86). FeLV-C emerges through a multi-tropic intermediate with point mutations that allow for infection via a heme export protein, FLVCR1 (87-90). This recurrent shift toward FLVCR1 usage is associated with red cell aplasia, perhaps a consequence of heme transport disruption during erythroid progenitor cell development. Porcine endogenous retroviruses are also divided into three groups with variable receptor usage, two of which are able to infect human cells and have thus been a major concern to the xenotransplantation field (91). PERV-A is able to infect human cells through the use of the gammahydroxybutyrate transport protein, PAR-1, or the riboflavin transport protein, PAR-2 (92). The PERV-B receptor is currently unknown.

Also in line with the recurring theme of transport proteins as gammaretrovirus receptors is the use of the sodium-dependent neutral amino acid transporter (ASCT2) by the baboon endogenous retrovirus (BaEV)/RD114 feline endogenous retrovirus/simian

retrovirus type D (SRV, including mason-pfizer monkey virus) interference group (93-95). BaEV, in addition to human ASCT2 and unlike the other members of its interference group, is also able to use ASCT1 from humans and mice (96). After the removal of protective N-linked glycosylations, mouse and hamster ASCT1 become functional receptors for the entire interference group (97). It should be noted that the usage of ASCT2 by SRV was likely made possible by an ancient recombination event between two quite distinct viruses. This hypothesis was prompted by the observation that reverse transcriptase phylogeny clearly classifies SRV as a betaretrovirus, while analyses using envelope sequences reveals close relatedness with gammaretroviruses in its interference group (i.e. BaEV and RD114) (98, 99). This emphasizes that virus classification that lacks inspection of receptor-binding regions can be misleading when used to predict receptor usage.

Sequence divergence in receptor binding regions does not necessarily predict that two viruses use distinct receptors. The HERV-W envelope SU is very distinct from that of RD114, BaEV, and SRV. Nevertheless, the ability of cells expressing ASCT2 to readily fuse with cells expressing the HERV-W envelope, syncytin, suggests that HERV-W also used ASCT2 as a cellular receptor during its exogenous replication (95, 100). The observation that syncytin is highly expressed in the placenta led to the hypothesis that its fusogenicity is responsible for syncytiotrophoblast formation during placental development. An *in vitro* model of cytotrophoblast maturation demonstrated that syncytin and ASCT2 expression co-linearly increase during differentiation and inhibition of syncytin expression leads to a decrease in cell fusions (101). Also in support of a cellular function for syncytin is the conservation of the following features among diverse

ape species and 24 human individuals: LTR promoter activity, splice sites required for an envelope mRNA, the envelope translational start site, and fusogenic activity (102).

Post-entry events and restriction factor encounters

TM-mediated fusion of the viral and cellular membranes allows the core of the virus to enter the cytoplasm of the target cell. Following entry, the capsid shell disassembles and the RNA genome is reverse transcribed to form a double-stranded DNA. During this time, the virus is susceptible to assault by host factors that are capable of capsid recognition, and others that act on the viral DNA during reverse transcription. Fv1 is a mouse gene that is similar to the mouse endogenous retrovirus-L Gag and is able to cofer strain-specific protection of mice from MLV infection (103-105). MLV variants displaying an N- and B-tropism, determined by differences in their capsid N-terminal domain, are resistant to the N and B alleles of Fv1, respectively (106, 107). Although the exact mechanism of restriction by Fv1 is elusive, it is clear that its capsid-specific action depends on Fv1 molecules in the target cell that are saturable at high multiplicities of infection. N-MLV is also sensitive to post-entry restriction by some tripartite motif (TRIM) 5α proteins (108-110).

TRIM5α is a primate restriction factor that was first recognized for its ability to form a major block to HIV-1 infection in rhesus macaques (108, 111, 112). Although human TRIM5α is unable to restrict HIV-1, it is active against N-tropic MLV (108-110). The species-specific sensitivities of viruses to TRIM5α, like Fv1 vulnerability, are determined by amino acid residues found in the capsid N-terminal domain that are exposed on the surface of the viral core prior to disassembly (113-115). TRIM proteins

share homologous RING finger, B-box, and coiled coil domains (116). In addition to this tripartite motif, TRIM5α also contains a SPRY/B30.2 domain. It has been demonstrated that the SPRY/B30.2 domain of TRIM5α is responsible for capsid recognition, while the RING and B-box domains contribute an unknown, but crucial function during restriction (117-121). Additionally, a coiled-coil domain allows TRIM5α multimerization and can allow non-functional truncated TRIMα mutants to have a dominant negative effect (121). The TRIM5α RING finger has been shown to have E3 ubiquitin ligase activity and induces auto-polyubiquitination, which can affect the half-life of the protein (122, 123). During restriction, TRIM5α has been observed to undergo proteosome-dependent degradation (124). The involvement of the proteosome in restriction is unclear, though, given that treatment of cells with proteosome inhibitors can cause an increase in the formation of reverse transcription products but are unable to cause an increase in infection during TRIM5α-mediated restriction (125-128).

A cellular prolyl isomerase is also able to modulate HIV-1 infection. It has been shown that cyclophilin A (CypA) interacts with HIV-1 Gag and is efficiently incorporates into virions (129-131). Disruption of CypA packaging with a small molecule (cyclosporine A) or mutation at a single proline residue in Gag decreased infection in human cell lines. In later work, though, expression of CypA in target rather than producer cells was shown to modulate HIV-1 infection and these effects differed according to both the capsid sequence and target cell line used (132, 133). For example, cyclosporin A treatment of Jurkat and HOS cells decreased infection. In contrast, treatment of HeLa and H9 cells did not affect infection with wildtype virus and actually enhanced infection with virus containing the A92E or G94D capsid cyclophilin A

resistance mutations (134, 135). Cyclophilin A in the target cell, therefore, has differential effects on infection depending on amino acid identities on the incoming capsid. Furthermore, the cell type specificity of these phenomena suggest a possible involvement of unknown host factors.

A specific linkage between cyclophilin A and restriction factors was demonstrated by the observation that disruption of its interaction with capsid using cyclosporine A or siRNA knockdown of target cell cyclosporine A was able to relieve the restriction of HIV-1 by rhesus macaque TRIM5α (136-138). Thus, it appears that the cyclophilin A-capsid interaction can be crucial to TRIM5α activity. In support of this assertion is the observation that convergent evolution has produced TRIM-cyclophilin A (TRIMCyp) fusion proteins in new world and old world monkeys by gene retrotransposition (139-144). Recurrent TRIMCyp formation emphasizes the antiviral potential that linkages between cyclophilin A and TRIM5 activities can offer and, furthermore, suggests that their combined function has provided primates with crucial protections from retroviruses throughout history. Additionally, high ratios of nonsynonymous to synonymous mutation have been observed in primate TRIM5α genes, suggesting that the evolution of these loci has been driven by positive selection (145, 146). The conclusion that TRIM5α genes have been under positive selection during primate evolution further suggests that they were rescuing primates from viral challenge long before the modern infections that we observe today.

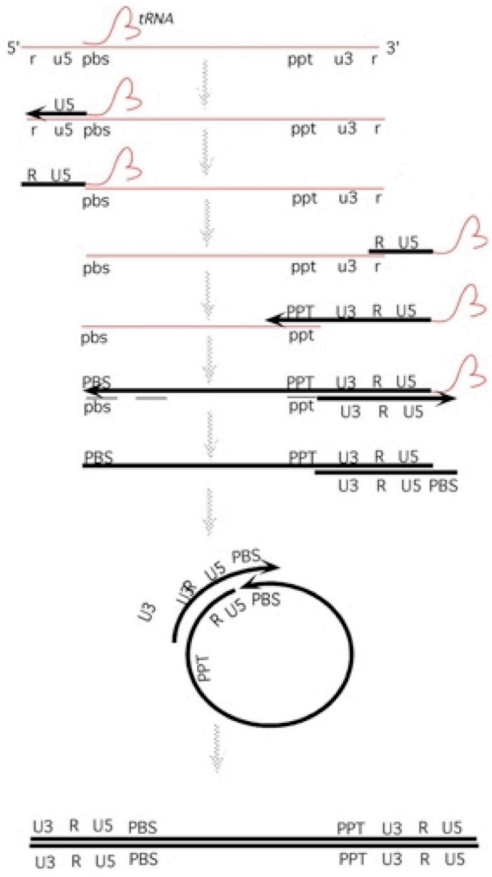


Figure 3. Reverse Transcription. RNA is represented by red lines and DNA, by thick black lines. Each step is described in the text. Drawing is by A. Telsnitsky and is taken from *Field's Virology,* Fifth Edition (147).

If the retroviral replication cycle is able to proceed normally, reverse transcriptase next synthesizes the viral cDNA to be integrated into the cellular genome (Figure 3) (148-150). DNA synthesis does not commence until after entry into the target cell, perhaps due to the lack of deoxyribonucleotides in the virion. Virion-packaged cellular tRNA is usurped for the initiation of reverse transcription at the primer-binding site (PBS), located just 3' of the 5' long terminal repeat (LTR). The 5' LTR is reverse transcribed first, forming the minus-strand strong-stop DNA before degradation of the

first template region by the RNase H domain of reverse transcriptase. The freed strong-stop DNA next binds to the opposite end of the viral RNA through complementation with sequence in its 3' end, thus allowing reverse transcription and the coupled RNase H activity to proceed through all open reading frames to the 5' end of the PBS. RNase H activity initially spares a polypurine tract (PPT) adjacent to the 3' LTR and this remaining RNA serves as a primer for synthesis of the 3'LTR positive strand before further RNA digestion. A new PBS is synthesized by the reverse transcription of the 5' end of the tRNA (plus-strand strong-stop DNA) prior to tRNA digestion, again by viral RNase H. Finally, the DNA-templated completion of both strands is prompted by circularization via base pairing at the PBS.

The existence of single-stranded DNA intermediates during reverse transcription leaves the virus vulnerable to attack by already present cytidine deaminases, carried from the virus-producing cell by the virion itself. The apolipoprotein B mRNA editing enzyme catalytic polypeptide (APOBEC) protein family of cytidine deaminases includes at least two restriction factors that are capable of acting potently on retroviruses- APOBEC3G and APOBEC3F (151-155). APOBEC3B, APOBEC3C, APOBEC3DE, and APOBEC3H have also been shown to have smaller degrees of antiviral activity (156-161). When APOBEC3 is expressed in virus-producing cells, they are incorporated into budding viral particles through their interaction with the nucleocapsid domain in an RNA-dependent manner (162-168). Following reverse transcription in the target cell, APOBEC3 proteins are able to deaminate deoxycytidine residues on the nascent negative sense cDNA to deoxyuridine (169-171). During synthesis of the DNA positive strand, the presence of aberrant deoxyuridine results in guanosine to adenosine mutation.

APOBEC3-induced cytidine deamination takes place with biases toward specific dinucleotide contexts. APOBEC3G prefers cytidines immediately 3' of another cytidine, which results in GG to AG mutations; while the other APOBEC3 proteins deaminate cytidine in a TC dinucleotide context and these events cause GA to AA mutations (APOBEC3G can also cause GA to AA mutations but far less often than GG to AG mutations) (153-155, 160, 161, 170, 171). The frequency of hypermutation can vary across the viral genome according to the length of time that single-stranded DNA is exposed prior to complementation to nascent positive-strand DNA (171). In HIV-1, this variability manifests as two gradients with increasing G to A mutation frequency in the 5' to 3' direction that peak just 5' of the polypurine tracts (172).

Human APOBEC3G was first identified as a restriction factor when it was found to determine the permissivity of cell lines to HIV-1 virions that lack the Vif protein (HIV-1 deltaVif) (152). Vif is able to inhibit APOBEC3G by recruiting it to a cullin-5/E2 ubiquitin ligase complex. This results in the polyubiquitination and subsequent proteosomal degradation of APOBEC3G before it can be incorporated into the budding virion (173, 174). The ability of Vif to inhibit APOBEC3 proteins varies in a species-specific manner (156). HIV-1 Vif is not able to induce the degradation of rhesus macaque APOBEC3G or APOBEC3F and this, in part, explains the inability of HIV to infect these animals (155, 175). The APOBEC3 gene family has expanded from one gene in mice to seven genes in primates and the primate genes, including APOBEC3F and APOBEC3G, have been under strong positive selection since before the divergence of old and new world monkeys 33 million years ago (176, 177). Therefore, it appears

that, like TRIM5 proteins, APOBEC3 provided a critical defense in primates against retroviruses long before the activities we observe today against modern viruses.

The restriction of Vif-deficient HIV-1 by APOBEC3F/G could take place through several possible mechanisms. First, the amount of proviral DNA is decreased during restriction (178, 179). This could be the result of the removal of uracil from deoxyuridine by uracil DNA glycosylase (UNG), followed by the action of cellular DNases. Indeed, UNG-2 is incorporated into virions but a depletion of UNG-2 in virus-producing cells is not able to rescue HIV-1 from restriction, leaving this mechanism unsupported (180). It has also been proposed that APOBEC3F- and APOBEC3G-mediated restriction may be independent of deaminase activity (181-184). APOBEC3F and APOBEC3G mutants containing changes to their C-terminal zinc coordinating catalytic site were not competent as deaminases but still retained at least 60% of their antiviral activity. Contrary to this, it had already been shown that mutating these same sites *does* destroy the antiviral activity of APOBEC3G (169). Furthermore, it has been argued that any deaminase-independent restriction observed is only made possible by *in vitro* over-expression of APOBEC3G, as evidenced by a complete lack of restriction capability in deaminase-deficient mutants when their expression levels are comparable to APOBEC3G expression in activated peripheral blood mononuclear cells (PBMC) (185-187). Although the *in vivo* applicability of a deaminase-independent restriction of HIV-1 by APOBEC3G remains controversial, a variety of mechanisms have been proposed including inhibition of tRNA priming, strand transfer, both early and late viral DNA production, and integration (188-191).

Perhaps the most intuitive explanation for the antiviral effects of A3 proteins, though, is a debilitation of viral progeny by G to A hypermutation in the provirus. This is especially likely in the case of APOBEC3G mediated restriction, which due to the bias in the dinucleotide context of APOBEC3G-triggered deamination, has a propensity to cause tryptophan codons (UGG) to mutate to stop codons (UAG). When both mutated dinucleotide contexts are considered, the high degree of hypermutation found in individual proviruses would likely impact their fitness through frequent nonsynonomous mutations.

Hypermutation in vivo

Investigations into the effects of hypermutation on HIV-1 are directly applicable to HIV evolution within an infected individual because, in spite of Vif activity, hypermutation does occur in vivo. Janini *et al.* analyzed protease genes integrated in peripheral blood mononuclear cells (PBMC) from patients at different disease stages and infected with a variety of HIV subtypes (192). Abundant hypermutated protease sequences were detected in 43% of the 53 patients studied. In the hypermutated sequences, 20 to 94% of G residues in the GG and GA dinucleotide contexts were mutated to A, which is consistent with APOBEC3G and APOBEC3F activity, respectively. Mutations in the GA context were twice as frequent as mutations in the GG context suggesting that APOBEC3 proteins other than APOBEC3G may predominate in HIV-1 hypermutation *in vivo*. Consistent with this hypothesis is the observation that HIV-1 Vif is less efficient at triggering APOBEC3F degradation compared to APOBEC3G degradation (155).

In another study, Kieffer *et al.* searched for hypermutated RT and protease sequences in both plasma virus and resting $CD4^+$T cells (193). Plasma virus and PBMC were isolated from nine patients and the PBMC was fractionated to isolate resting $CD4^+$T cells. The patients had a undetectable viremia due to suppression by highly active antiretroviral therapy (HAART). Resting $CD4^+$T cells were studied because even when viremia is kept at low levels by HAART, integrated latent viral proviruses remain abundant in these cells. There were no hypermutated sequences detected in plasma virus; but in the case of resting $CD4^+$T cells, at least one hypermutated sequence was found in each patient. The hypermutation was manifested as G to A changes with 98% of these changes occurring in the GA or GG dinucleotide context. Unlike the study conducted by Janini *et al.*, mutations within the GG context were four times more common than mutations within the GA context. Perhaps the use of resting $CD4^+$T cells, as opposed to unfractionated PBMC, from patients on successful HAART narrows the study to a latent pool containing inactive hypermutated proviruses that have primarily been mutated by APOBEC3G. In support of this assertion is the observation that APOBEC3G is more active against HIV delta Vif than APOBEC3F and has a greater propensity to induce stop codons (155). Additionally, APOBEC3G has been observed in cell culture experiments to be capable of causing a small amount of GG to AG mutations in HIV sequences, even in the presence of Vif (194).

Overall, it is unclear which APOBEC3 protein(s) are predominately active against HIV *in vivo*. The preferred dinucleotide context of G to A mutations may vary greatly within a single patient. It is likely that the ability of each deaminase to hypermutate HIV depends on the potency of the Vif species encountered. It has been shown that a variety

of single amino acid changes in Vif can affect its specificity for APOBEC3F, APOBEC3G or both (195, 196). A study by Simon *et al.* described highly variable Vif sequences within individual patients and inactive Vif genes were isolated from all seven patients studied (197). Furthermore, Vif genes were isolated that were active against APOBEC3F and APOBEC3G, APOBEC3F but not APOBEC3G, or vice versa. Concordantly, the predominance of either GG to AG or GA to AA substitutions varied between hypermutated proviruses. This suggests that Vif variability can allow for a large range in both the quantity and the dinucleotide context of hypermutation *in vivo*.

Aside from the generation of stop codons by APOBEC3G activity, any increase in the amount of random mutation could be very deleterious to the overall fitness of an HIV-1 population. It has been argued that HIV-1 reverse transcriptase has evolved a fidelity high enough to maintain sufficient sequence integrity, but low enough to allow rapid sequence diversification (198). An increase in the frequency of random mutations may push the mutation rate high enough to allow the accumulation of a significant number of deleterious mutations, thus giving rise to a genetic load with significant impacts on the overall population fitness.

On the other hand, the ability of HIV reverse transcriptase to generate recombinant genomes could allow for the continual regeneration of fit viruses despite the generation of hypermutants. Recombination in retroviruses is possible because it is an inherent trait of retroviral virions to contain two viral genomes (199-201). When virus is produced from a cell containing multiple infections, packaging of homologous RNA from more than one proviral source can occur. This results in the presence of differing templates for reverse transcription during the subsequent infection. It appears that

reverse transcriptase is capable of jumping between templates multiple times during a single cycle of replication (202, 203). The resulting viral DNA that continues on to integration is, thus, a recombinant of the two genomes used during reverse transcription. It has long been theorized that recombination has repeatedly arisen throughout biology, in part, to ameliorate the effects that a load of deleterious mutations can have on a population (204-207). It is possible that it is the high frequency of recombination in HIV-1 that allows it to cope with high mutation rates and, perhaps, even make use of the variability that hypermutation offers. Recombination between hypermutants and genomes untouched by deamination may generate fit viruses containing a relatively low-level of APOBEC3-induced mutation.

Mutations that resemble those imposed by APOBEC3 proteins (i.e. G to A within GG or GA contexts) have been observed in viruses that have a low burden of mutation compared to proviruses that have been crippled by hypermutation (208-210). Low levels of putative APOBEC3-induced mutation were shown to include residues within confirmed cytotoxic T lymphocyte epitopes that evolved rapidly during early infection (210). This suggests that variability generated by APOBEC3 activity can be utilized by HIV-1 during immune evasion. Furthermore, the occurrence of drug resistance mutations within sequence contexts targeted by APOBEC3 suggests that hypermutation also has the potential to contribute to drug resistance evolution (reviewed in Chapter IV). APOBEC3 activity *in vitro* has been shown to induce the emergence of 3TC escape mutants, though this required the cloning of hypermutated loci from unfit variants into wildtype virus (211). It is possible that this recapitulated recombination events that can occur between replication competent virus and inactivated hypermutants *in vivo*. So, the question arises

for a viral population within an infected host: is hypermutation induced by APOBEC3 a burden or a benefit?

Post-reverse transcription events and production of viral progeny

Following reverse transcription, the newly synthesized viral DNA transports to the nucleus as part of a pre-integration complex that includes integrase, reverse transcriptase, and Gag components (212, 213). The biochemical steps of integration have been extensively studied *in vitro* [reviewed in (214) and Figure 4]. First, integrase transforms the blunt ends of the provirus into 5' overhangs by the removal of two nucleotides, typically TT, leaving recessed strands ending in CA. Second, the newly formed 3' hydroxyls each undergo a nucleophilic attack of phosphorous in the cellular DNA. The resulting transesterification generates single-stranded nicks between the 5' end of each proviral DNA strand and the displaced 3' ends of the cellular DNA. Two nucleotides at the free 5' ends of viral DNA and several cellular nucleotides on the opposite, ligated strand lack base pairing. These structures are promptly repaired, likely by cellular machinery, with the insertion of nucleotides that are complementary to unpaired cellular sequence followed by a final transesterification that replaces the two unpaired viral nucleotides and removes the nick.

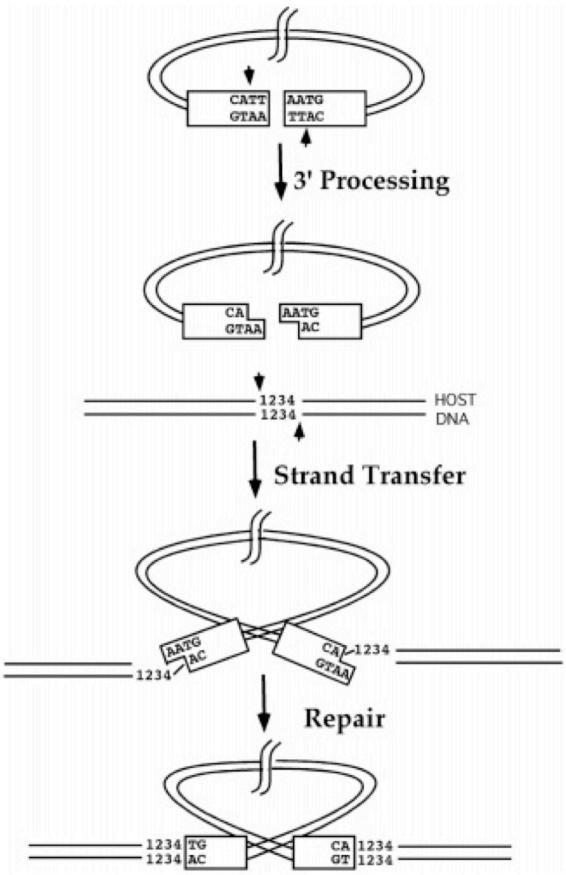


Figure 4. Retroviral integration. The steps involved in the formation of provirus are shown. These include processing of the viral DNA ends by integrase, nucleophilic attack of cellular DNA phosphorous atoms by 3' hydroxyls on the viral DNA (or 'strand transfer,' also catalyzed by integrase), and finally repair single-stranded nicks. Adapted from Goff, *Ann. Rev. Gen.* 1992 *(214)*.

With the aid of cellular factors, the integrated provirus is now available for both gene expression and inheritance by cellular progeny. The integrated viral LTR contains promoter elements that prompt transcription by the cellular RNA polymerase (215-217). In the case of HIV-1, the viral Tat protein stimulates transcription through the association of a cyclin T/CDK9 complex with the viral TAR RNA nascent transcript (218-228). This results in hyperphosphorylation of the RNA polymerase II C-terminal domain and,

consequently, transcript elongation (225-228). Full-length viral transcripts are capped, poly-adenylated, and exported for either virion packaging or translation of Gag and GagPol. Other transcripts are spliced for envelope translation and during the production of complex retroviruses, alternative splicing allows for the expression of several accessory genes (229). In HIV-1, the export of intron-containing transcripts requires the binding of the virus-encoded Rev protein to Rev-responsive elements in the RNA (230-236).

Translation of unspliced mRNA produces both Gag and GagPol proteins. The inclusion of Pol on 5-10% of Gag proteins occurs when structured RNA components in the transcript cause the ribosome to "ignore" the Gag stop codon. In some retroviruses (e.g. gammaretroviruses) a tRNA is utilized at the stop codon, while in other viruses (e.g. lentiviruses) slippage of the ribosome causes a frame shift (237-239). The Gag and GagPol proteins accumulate on viral RNA at the plasma membrane to form nascent virions (240-242). Fission of viral and plasma membrane allows assembled particles to bud from the cell surface. The membrane manipulations required for such an event are catalyzed by cellular class E vacuolar sorting machinery, which are proteins that normally execute the topologically similar formation of multi-vesicular bodies [reviewed in (243, 244)]. Recruitment of class E machinery requires Gag amino acid motifs, termed late-budding domains (PT/SAP, LYPxL, or PPxY). Finally, processing of Gag and GagPol by the viral protease facilitates maturation of the budded particle (245, 246). Gag and GagPol cleavages produce mature matrix, capsid, nucleocapsid, protease, reverse transcriptase, and integrase proteins.

Following the budding of a new virion, and before escape from the surface of the very cell that produced it, the virus may be assaulted by the cell through the action of the tetherin protein. Tetherin was identified as a host restriction factor based on both its interferon induction and specific expression in cells that were non-permissive to replication of HIV-1 lacking expression of the Vpu accessory protein (247, 248). Biochemical characterization and the replacement of tetherin components with analogous protein structures have demonstrated that tetherin accomplishes restriction through a direct bridging of virions to cells, with one end of tetherin attached to the viral membrane and the other to the plasma membrane (249). Importantly, in the absence of specific antagonism, tetherin seems to act broadly against enveloped viruses without any identified protein specificity (250).

Endogenous retroviruses and the lessons that they can offer

We refer to the cycle of extracellular virion production and subsequent cellular infection as exogenous virus replication. There is utility in such a designation because the mere presence of provirus in cellular genomes also allows for its replication by the cell's indiscriminate DNA polymerase. Such endogenous replication dictates that any descendent of an infected cell will contain a copy of the integrated provirus. Robin Weiss first pointed out that heritable elements in cells from chickens were capable of allowing the replication of Rous sarcoma viruses (RSV) that were normally incompetent without the aid of a helper virus (5). Later Weiss and Peter Vogt were able to use carcinogens to induce the release of avian tumor viruses from uninfected chicken cells (251). These pivotal observations, in combination with the observed heritability of some

avian leukosis virus (ALV) antigens (252), led to the theory that retroviruses are capable of invading the host germ line and can consequently become endogenized into the organism-wide genome of progeny.

Sequencing of primate genomes has expanded our appreciation for retrovirus endogenizations that have taken place during our own evolution. At least 8% of human genomic sequence is the result of retroviral infections (253). Genomic sequence, therefore, provides an ancient historical account of past retroviral epidemics. Experiments with human endogenous retrovirus K (HERV-K), a virus that was repeatedly endogenized in the ancestors of modern primates for millions of years, demonstrated that our understanding of ancient viruses does not have to end with their sequencing (13, 254). A consensus of inactive human-specific HERV-K integrations was constructed and gave rise to infectious virus particles *in vitro*. The virus, perhaps not produced for over one million years, was further characterized as being resistant to TRIM5 and APOBEC3G restriction factors, but sensitive to human APOBEC3F. It may, therefore, be the case that members of the HERV-K family contributed to the positive selection imposed on the APOBEC3F locus over the course of its evolution. These initial endeavors in "paleovirology" began to demonstrate the potential that such studies have in describing the long-term co-evolution of hosts and retroviruses.

The work described herein, sought to extend the exploration of ancient viral challenges to include primate gammaretroviruses. A comparison of the chimpanzee genome with our own revealed that there were distinct differences in our retroviral encounters since the divergence from a common ancestor six million years ago (8, 9). In fact, 7% of chimpanzee-human INDEL variation is attributable to retrovirus integrations.

A close relative of the modern MLV, CERV2, and a more distant relative, CERV1, are present in the genomes of chimpanzees (Figure 5) and other non-human primates, but are entirely absent from the human genome (8, 9, 20). Therefore, CERV1 and CERV2 may reveal genetic factors that have influenced cross-species transmissions over the course of chimpanzee and human speciation. Furthermore, such investigations could support a notion that a broad repertoire of retroviruses has imposed selective pressures that have shaped the evolution of modern host factors.

First, we asked whether host factors that restrict modern retroviruses were able to target CERV1 and CERV2 during their exogenous replication (255). Due to the close relatedness of CERV1 and CERV2 with moloney MLV, we were able to construct functional recombinants containing CERV1 or CERV2 capsid N-terminal domains within an MLV GagPol context. The resulting chimeric virions were shown to be largely resistant to restriction by human, chimpanzee, and rhesus macaque TRIM5α. Therefore, we cannot speculate that TRIM5α was involved in a protection of human ancestors from CERV1 or CERV2 infection. An inspection of CERV1 and CERV2 sequences, though, revealed scars likely made by APOBEC3G. Namely, G-to-A mutations were observed with bias toward GG dinucleotides. Homologues in the rhesus macaque carried G to A mutations with both GG and GA dinucleotide contexts implying encounters with multiple APOBEC3 proteins. Further work *in vitro* demonstrated that APOBEC3G is capable of restricting infection by virions carrying CERV2 structural proteins. APOBEC3 proteins were therefore likely to be capable of restricting these viruses and, in return, infections by them may have contributed to the positive selection imposed on APOBEC3. Furthermore, debilitating APOBEC3-induced mutation is likely responsible for the

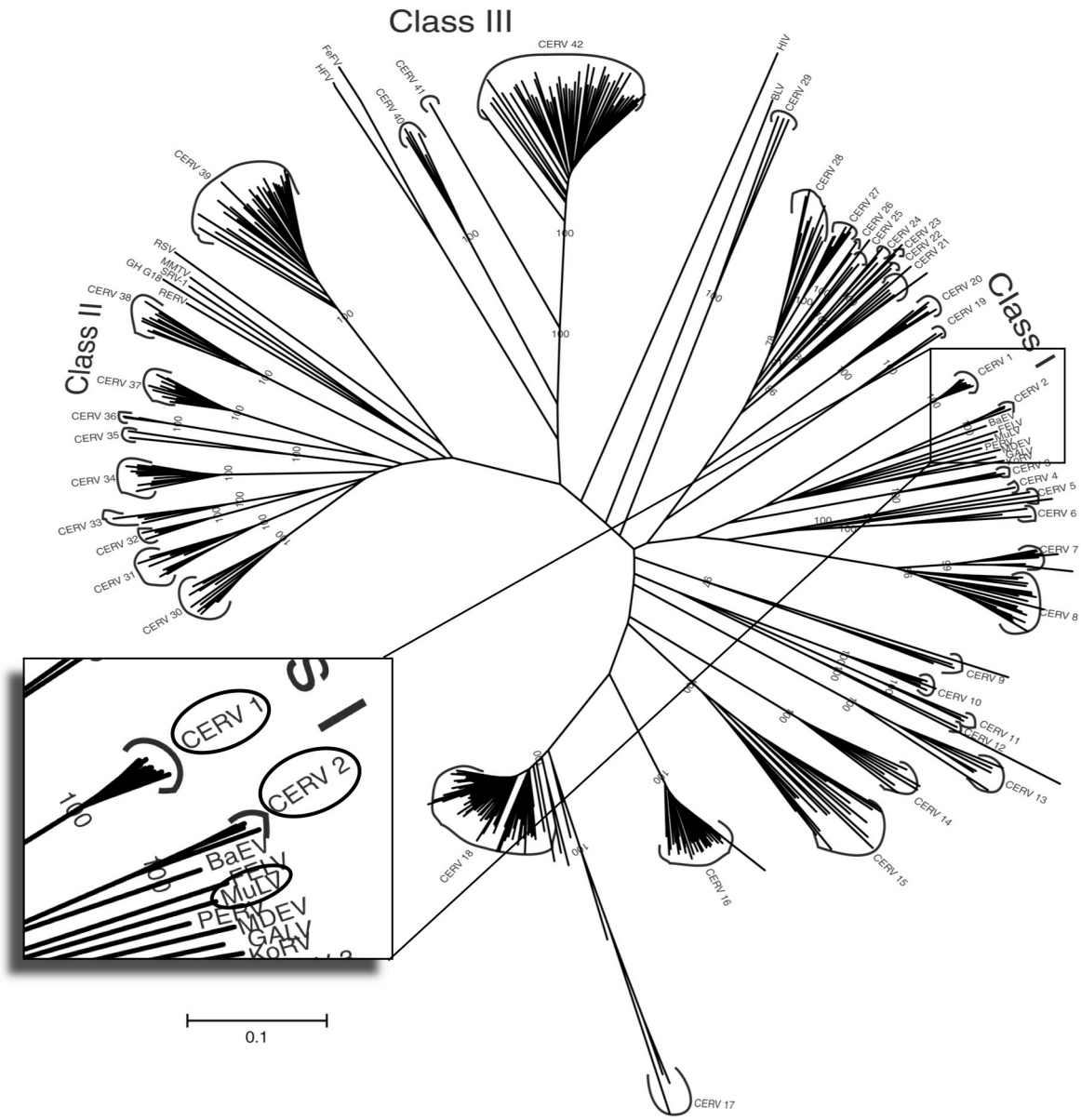


Figure 5. Phylogeny of chimpanzee endogenous retroviruses (CERVs). The unrooted neighbor joining tree is based on CERV reverse transcriptase sequences and those from several representative retroviruses as indicated. Three classes of endogenous retroviruses are designated: class I is related to modern gammaretroviruses, class II is related to modern betaretroviruses (class II CERVs are homologous to HERV-K), and class III is distantly related to modern spumaviruses. An enlarged portion is shown to indicate the position of CERV1, CERV2, and MLV on the tree. Adapted from Polavarapu *et al. Genome Biology* 2006 (8).

inactivity of many of the individual proviruses. This work allows us to speculate about a role of APOBEC3 proteins in limiting exogenous replication of these viruses.

The potential impacts that APOBEC3-induced mutation can have on an evolving population were studied using HIV-1 *in vitro*. Virus was allowed to replicate in the presence of varying degrees of APOBEC3 activity. It was evaluated whether the resulting hypermutation was capable of accelerating drug resistance evolution or if it constitutes a genetic load that hinders the emergence of escape mutants. Populations of HIV-1 delta-Vif that carried significant genetic loads were not observed to acquire drug resistance, even when challenged in fully permissive cells. A low level of APOBEC3-associated mutation was also observed in Vif-positive virus and this modest degree of diversity did not impede and may have contributed to its evolution.

To extend our understanding of host factors that did not assault, but rather aided CERV2 during the time that it replicated, we constructed a functional consensus envelope gene. MLV particles carrying this envelope displayed a broad species tropism in tissue culture and were further used in a genetic screen that identified copper transport protein 1 (CTR-1) as a novel cell-surface receptor that was used by CERV2 during ancient times. The only species tested that was non-permissive to CERV2 was the hamster and its lack of permissivity can be explained by subtle differences in CTR1. The identity of the CERV2 receptor does not predict a limited primate host range. This work does demonstrate, though, that reconstructions of ancient envelope genes can allow the discovery of host proteins that have been used as viral receptors throughout evolutionary history.

Chapter II. Materials and Methods

Virions and cell lines

Production of virions was accomplished by polyethyleneimine (PEI) cotransfection of 293T cells. A three-plasmid system was used to generate virus particles that were only capable of a single cycle of replication. Two scales are described for production in either 24-well or 100mm dishes. The transfections included plasmids expressing moloney murine leukemia virus GagPol or HIV-1 GagPol from the NL4.3 laboratory-adapted strain (500ng or 6μg); a vector containing an MLV or HIV-1 packaging sequence proceeding a reporter gene and/or a selectable marker (500ng or 6μg); and an envelope glycoprotein from either vesicular stomatitis virus (VSV) for its high infectious titer in a broad range of target cells (100ng or 1μg), or a retroviral envelope as indicated in the text (500ng or 6 μg). The reporter vectors used were CNCG (256) and CSGW (257), which are MLV and HIV vectors, respectively, that both contain an eGFP expression cassette.

Plasmids were mixed in serum-free DMEM (50μL or 1mL) followed by addition of 4μg PEI per 1μg of total DNA (from a 1μg/μL PEI stock solution). This mixture was vortexed briefly and incubated at room temperature for 10 minutes before addition to 293T cells, which were 60-80% confluent in freshly added media (500μL or 5mL). Transfections were allowed to proceed overnight before a change to fresh media (500μL or 8mL). Supernatants were collected 48 hours post-transfection and purified through 0.22μm filters. All infections were allowed to proceed overnight and 5μg/mL polybrene

was included on all cells except CHO-PGSA. A detailed description of MLV particles containing endogenous capsid N-terminal domains can be found in a separate section below.

The LMN8iresZeo cDNA plasmid library was a gift from Stuart Neil and was generated as follows. Total RNA was extracted using Invitrogen TRIzol from $5x10^6$ HeLa, purified with two sequential chloroform extractions, precipitated in isopropanol, and washed in 70% ethanol. mRNA was purified from total RNA using Qiagen Oligotex polyA+ resin. cDNA containing SfiI sites was synthesized and amplified using Clontech SMART primers and Invitrogen size exclusion columns were used to dispose of small fragments. Following precipitation and SfiI digestion, cDNA was ligated overnight to LMN8iresZeo (an MLV vector with a zeocin resistance gene translated via an internal ribosomal entry site). To assess library complexity, 10% of the ligation was transformed and serial dilutions were plated. The remaining ligation was transformed and grown under selection in soft agar for 72 hours.

Virus stocks used to transduce CHO-PGSA cells with the HeLa cDNA library were titrated by infection of 10^5 CHO-PGSA cells per well on 12-well plates with a 5-fold serial dilution of virus. The following day, infected cells were re-plated on 100mm dishes and 200μg/mL zeocin was added that evening. Cell foci were counted 10 days later to determine the number of infectious units per unit volume of the virus stock solution (IU/mL). An MLV vector expressing a neomycin resistance gene under a CMV promoter (CN) was used to generate CERV2 enveloped particles for the first round of library screening. CN was constucted by the removal of CMV-DsRED from CNCR using BstBI/NotI, followed by klenow treatment and blunt-ended ligation. CERV2

enveloped virions carrying CN or LHCX (a Clontech MLV vector expressing a hygromycin resistance gene) were titrated on HeLa cells using 1μg/mL G418 or 100μg/mL hygromycin, respectively, in a fashion similar to the cDNA library titration.

CHO, PGSA, and all derivatives were cultured in ham's F12 media supplemented with 10% fetal bovine serum and 1mM glutamine. *Pan troglodytes verus* skin fibroblasts AG06939 (Coriell) were grown in MEM alpha with 20% fetal bovine serum and 2mM glutamine. MT2 and CEMSS cells were grown in RPMI with 10% fetal bovine serum and 2mM glutamine. All other cell lines were grown in DMEM supplemented with 10% fetal bovine serum. All media was supplemented with 10μg/mL gentamycin. The CHO-CTR1 and PGSA-CTR1 cell lines were generated by cloning CTR1 into LNCX2 (a Clontech MLV vector containing an LTR-driven neomycin resistance gene and a multiple-cloning site proceeded 3' of a CMV promoter) or LNCX2 constructs containing an HA tag before or after the multiple-cloning site using XhoI and NotI (forward primer: 5'- TATATACTCGagATGGATCATTCCCACCATATGG -3' reverse primer: 5'-TATATAGCGGCCGCTCAATGGCAATGCTCTGTGATATCC -3'). In the case of CTR1 mutant constructs, PCR products were first cloned using XhoI and NotI into PCR3.1 (Invitrogen) containing an HA tag sequence 5' to the cloning site and then subcloned to LNCX2 using SnaBI and NotI. Hamster CTR1 was amplified from CHO cDNA using primers designed against untranslated regions that are conserved between mouse and rat CTR1 (forward primer: 5'-TATATACTCGagGGGATCCAGTTCTGagAGGAAGAC -3' reverse primer 5'-TATATAGCGGCCGCAGACAGGagAGagRAGGGATGATTGG -3'). Following XhoI/NotI cloning into pCR3.1 and sequencing, primers were designed to amplify only

the open reading frame of hamster CTR1 (forward primer: 5'-TATATACTCGagATGGGGATGAACCATATGGGC -3' reverse primer 5'-TATATAGCGGCCGCCTTAATGGCAATGCTCTGTGATGTCC -3') and this product was cloned into LNCX2 using XhoI and NotI. VSV-G enveloped MLV pseudotypes were used to transduced CHO or PGSA cells prior to selection with 1mg/mL G418.

Full-length HIV-1

A pNL4.3 HIV-1 provirus that lacked Vpr expression due to both a start codon mutation and the introduction of a stop codon (both silent in the Vif open reading frame) was provided by Stuart Neil. Vif expression was disrupted in this construct by the introduction of both a 13-nucleotide deletion and an in-frame stop codon to the region immediately following the end of the GagPol open reading frame. This was done by the amplification of a region from the AgeI restriction site in Pol to the border of the Vif deletion and a second region from the end of the Vif deletion to the EcoRI restriction site in Vpr. The primers were designed to generate products that contained a 29 nucleotide overlapping sequence that included the introduced Vif mutations (sense primer: 5'-GTAAAACACCATATGAAGCTTAGGACTGGTTTTATAGAC -3' and antisense primer: 5'- CCAGTCCTAAGCTTCATATGGTGTTTTACTAATCTTTTCCATGTG -3'). Following an overlapping PCR that included the products of the first PCRs as templates and external primers, the final PCR product was gel-purified and cloned into the parental pNL4.3 delta Vpr using AgeI and EcoRI. Next, in order to acquire an eGFP reporter, SalI and SpeI were used to sub-clone both the deltaVpr and deltaVpr/deltaVif genotypes into an HIV-1 proviral plasmid that expressed NL4.3 GagPol, HXB-2

envelope, and eGFP in place of Nef (previously derived from pNL/HXB and R7/3/GFP (258, 259)). Virus stocks were generated by PEI transfection of 293T cells on 100mm dishes using 5μg of proviral plasmids. The resulting viruses are referred to in the text as Vif+ and ΔVif.

Endogenous capsid sequence retrieval

To identify MLV-related viruses in primate genomes, the amino acid sequence of the capsid N-terminal domain of moloney murine leukemia virus was used in a TBLASTN search of the chimpanzee genome (http://www.ensembl.org). After retrieving all complete TBLASTN hits, the sequences were separated into two families, corresponding to CERV1 and CERV2, according to sequence homology and length. After removing redundant sequences, ClustalW was used to obtain nucleotide and amino acid consensus sequences for both CERV1 and CERV2 capsid N-terminal domains (CA-NTDs). The same method was used to find homologues of CERV1 and CERV2 CA-NTDs in *Macaca mulatta* (RhERV1 and RhERV2, respectively). RhERV1 CA-NTDs included two phylogenetically distinct subgroups, termed RhERV1a and RhERV1b, and two independent consensus sequences were derived. Similarly, CA-NTD sequences from enMLV were retrieved from the C57BL/6J genomic database. A phylogeny was reconstructed for all capsid NTDs using ClustalX1.8 and phylogenetic trees were drawn using Figtree.

Individual CA-NTD amino acid sequences were used in TBLASTN searches and regions positioned immediately 3' to precise hits were inspected for Env-like sequences. The moloney MLV was used to assist in the definition of a theoretical

CERV1/CERV2/RhERV1/RhERV2 open reading frames. To define a consensus envelope sequence for each endogenous retrovirus family, a single envelope amino acid sequence was used as a TBLASTN query. All resulting envelope sequences were aligned to derive a consensus sequence, regardless of their linkage to a previously retrieved capsid sequence. Analyses of the CA-linked envelope genes were then performed as for the CA-NTD sequences. In the case of RhERV2, two distinct groups of envelope sequences were observed and separate alignments and consensus sequences were derived for each subgroup of envelopes. Similarly, three distinct enMLV envelopes were grouped according to the previously defined tropisms of these viruses: polytropic, modified polytropic, and xenotropic (72). In the case of CERV1, because of the large numbers of unique CA-NTD sequences, the analysis was confined to two groups of 10 CA-NTDs that were each used to define linked Env sequences. One CA-NTD group had zero or one G to A mutations, and a second group had ≥4 G to A mutations.

Because of the small number of CERV2 integrations in the chimpanzee germ line, all CERV2 proviral sequences were compiled. Flanking regions 1500 nucleotides 5' and 8000 nucleotides 3' to the capsid NTDs were aligned. Four of the proviruses aligned for 8293 nucleotides with pairwise identity scores of at least 83%. Therefore, those portions of sequence were trimmed and defined as full length CERV2 provirus. Six additional incomplete proviruses were defined. One database derived proviral sequence was excluded because a 99.9% sequence identity with another locus made it unlikely to be an independent retroviral integration, and is instead likely to be either a genomic duplication or an error in the compiling of genomic sequence.

Generation of endogenous retrovirus CA-NTD/MLV chimera libraries

A library of chimeric GagPol plasmids was generated that contained MoMLV GagPol with the capsid NTD replaced with that from endogenous retroviruses. Genomic DNA was isolated using standard protocols from *Pan troglodytes verus* skin fibroblasts AG06939 (Coriell), *Macaca mulatta* 221 cells, and NIH 3T3 cells and used as PCR template. Primers were designed to anneal to the capsid NTD consensus sequences described above. Because of relatively high conservation at the 3' end of the capsid NTD, the same reverse primer was used for all endogenous retrovirus families: 5'-TACYTTRGCCAAATTRGTRGG -3'. The 5' primers were specific to each virus family and all contained a BsmI restriction site: CERV1: 5'-CTCGCAGGCATTCCCCCTTCGGGAAATAGG -3'; CERV2: 5'-CTCGCAGGCATTCCCCCTCCGCACCGTG -3'. To amplify capsid genes from the Rhesus macaque and mouse genomes, primers were designed to anneal to a p12 region immediately 5' of the capsid- RERV1a: 5'-CAGCYGCCTGACTCYAYGGTGGCATTCCCCCTT -3'; RERV1b: 5'-CRACTCCCTGACTCCACYGTGGCATTCCCTCTC -3'; RERV2: 5'-CCTTCYACTTGGCAATCCTCGGCATTCCCCCTC -3'; and enMLV: 5'-GCRGAYTCCACCWCCTCYCRGGCATTCCCACTC -3'. A small fragment was amplified from MLV GagPol using the reverse complement of the capsid NTD 3' primer: 5'- CCYACYAATTTGGCYAARGTA -3'; and a primer annealing in the MLV capsid C-terminal domain: 5'- CTTCTAACCTCTCTAACTTTCTCC -3'. The amplified portion of the MLV GagPol plasmid was engineered to contain an AfeI restriction site. Thereafter, an overlapping PCR product was generated that included the amplified

endogenous virus capsid NTDs and the MoMLV capsid CTD fragment. This PCR product was cloned into the MLV GagPol plasmid using BsmI and AfeI restriction enzymes. The resulting chimeric GagPol plasmids were isolated and screened for functional capsid genes by cotransfection with an MLV-based vector containing GFP and VSV-G envelope using polyethyleneimine (PEI) in 293T cells. Supernatant from these cells were used to infect hamster CHO-KI-derived cells. Two days post-infection, cells were trypsinized, fixed in 2% PFA and subjected to FACS analysis using a Guava Easycyte to determine the percentage infected (GFP-positive) cells. To measure TRIM5α, TRIMCyp and Fv1 sensitivity, Mus dunni tail fibroblasts (MDTF) or MDTF stably expressing human TRIM5α, chimpanzee TRIM5α, rhesus macaque TRIM5α, African green monkey TRIM5α, owl monkey TRIMCyp, $Fv1^{N}$, $Fv1^{B}$, or $Fv1^{NB}$ that have previously been described (108, 121) were infected in the presence of 5μg/mL polybrene. GFP-positive cells were quantified by FACS analysis two days post-infection.

To determine whether the E91K mutation in the RhERV2 CA-NTD modulates its sensitivity to human and chimpanzee TRIM5α, this point mutation was introduced by PCR into a chimeric MLV GagPol plasmid containing a consensus RhERV2 CA-NTD sequence. Additionally, to determine whether the CA-CTD altered CERV1 sensitivity to TRIM5α, a CERV1 CTD sequence, as described in Kaiser S.M. *et al.* (260), was synthesized using as series of overlapping ~60 nucleotide oligonucleotides. The resulting product was then used in an overlapping PCR reaction with the CERV1 CA-NTD consensus sequence and this full-length capsid was cloned into the MLV GagPol plasmid. In order to match the CERV1 CA-NTD consensus to the ancestral sequence

described in Kaiser *et al*.(260), the R35Q mutation was also introduced using PCR-based mutagenesis into this CERV1 full-length capsid/MLV chimera.

Sequence analysis

Mutations and their dinucleotide sequence contexts were quantified using Hypermut (http://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html). We generated fasta alignments using MacVector to use as input for Hypermut and the consensus sequence for each virus family (described above) was used as a reference sequence. The Hypermut output provided the total number nucleotide changes (each insertion/deletion was reduced to a single nucleotide change in MacVector) as well as the number of G to A mutations relative to the consensus sequence. Hypermut also provided the number of G to A changes that occurred in each of the four possible dinucleotide contexts when the nucleotide immediately 3' to the G (or the +1 position) in the consensus sequence is considered (i.e. GG, GA, GC and GT). To quantify the -1 position for C to T changes, the +1 position for C to T changes, and the -1 position for G to A changes, the reverse complement, complement, and reverse of each alignment was generated respectively and used as Hypermut inputs. The latter two analyses allowed us to quantify the number of C to T and G to A changes occurring in CG dinucleotides. We repeated this analysis, after removal of G to A and C to T changes occurring in CG dinucleotides. In some charts, the percentages of G to A or C to T changes that were in each dinucleotide context were normalized according to the dinucleotide composition of a given consensus sequence. P-values were calculated for these data using a chi-square goodness of fit statistical test.

Gene Synthesis

ClustalW alignments and majority consensus sequences of CERV2 Gag, CERV1 envelope, CERV2 envelope, and RhERV2 envelope were derived using MacVector. All envelope sequences were collected from Ensembl using TBLASTN according to their homology with MoMLV protein sequences first and then CERV consensus sequences were used to blast the rhesus macaque genome to ensure the recovery of all homologues. Sequences flanking the CA-NTDs used in the hypermutation analysis described above were aligned with MoMLV Gag to define a CERV2 Gag open reading frame. As described in the next chapter, a second version of CERV2 Gag was also designed that included protease cleavage sites that matched those from MLV Gag (CERV2/MLVGag). Overlapping oligonucleotides were designed using Genedesign from Johns Hopkins University (http://slam.bs.jhmi.edu/gd). The target oligonucleotide length and overlap melting temperature was set to 60 nucleotides and 56°C, respectively. EcoRI and NotI sites were added to oligonucleotides designed to construct the 5' and 3' ends of the envelope genes for cloning into the PCAGGS vector (BCCM). In the case of CERV2 Gag, the 5' end was also designed for EcoRI recognition and at the 3' end, 20 nucleotides were designed to overlap with the 5' end of MLV Pol (forward primer: 5'-GagGagGAATTCATGGGACAGACTCTGACGAC -3' reverse primer: 5'-CTCCTGACCCTGACCTCCCTAATCGCTGTCCTCTCCCAGGGTCAGGACCGGGGTGGGTTTCTTCC -3'). Primers used to amplify MLV Pol introduced 20-nucleotides at the 5' end that overlapped with the 3' end of CERV2 Gag and a NotI restriction site was introduced at the 3' end (forward primer: 5'-

CCTGGGagAGGACAGCGATTAGGGagGTCAGGGTCAGGag -3' reverse primer: 5'-TATATAGCGGCCGCTTAGGGGGCCTCGCGGGTTAACC -3').

For envelope gene construction, two sequential polymerase chain reactions were carried out. The first contained a mixture of the overlapping oligonucleotides prepared by combining 1µL of each 100µM stock solution and then diluting ten-fold. A 50µL PCR reaction contained 3µL of the diluted oligonucleotide mixture and gene synthesis was carried out for 20 cycles using Phinzymes Phusion (1 min. 98°C, 1 min. 50°C, and 2 min. 72°C). The second PCR contained 2µL of crude product from the first in a 50µL reaction with 200nM external primers. Amplification proceeded for 25 cycles (1 min. 98°C, 1 min. 55°C, and 2 min. 72°C) and the resulting product was gel-purified using the Qiagen gel extraction kit and cloned into PCAGGS with EcoRI and NotI. Sub-cloning between individual clones was necessary to eliminate errors in CERV1 envelope using SpeI and RhERV2 envelope using DraI.

To construct the CERV2 Gag consensus fused to MLV Pol, first, the Gag was synthesized much the same way as was described above for envelope genes except, due to lower yields, there were a few protocol changes. The PCR buffer used in a 20-cyle assembly PCR was supplemented with 3mM $MgCl_2$ and then 3µL of this PCR was subjected to 15 cycles of amplification using external primers. A 1.8kb product was gel-purified for use in a third PCR that ran for 20 cycles. The resulting product was gel-purified and used in an overlapping PCR with MLV Pol. The two templates were used at equal ratio in 3% DMSO for a Phusion catalyzed, twenty-cycle reaction (1 min. 98°C, 1 min. 70°C, and 1.5 min. 72°C). After a final Qiagen gel-purification, the CERV2Gag-MLVPol PCR product was cloned into PCAGGS using EcoRI and NotI. To eliminate

errors, sub-cloning was completed using KasI and NheI for CERV2 Gag and XhoI for CERV2/MLV Gag.

APOBEC restriction assays

Human and rhesus macaque APOBEC3G with an N-terminal Myc tag was expressed from previously described pCR3.1 constructs (156). For comparative purposes, chimpanzee APOBEC3G was amplified from a pcDNA3.1 construct provided by Nathaniel Landau (261) and used to replace human APOBEC3G in pCR3.1-Myc using EcoRI and NotI. All other human APOBECs were expressed with C-terminal HA tags from previously described pCMV4 constructs (160). Virions were produced in the presence or absence of APOBEC proteins by cotransfection of 293T cells on 24-well plates with 300ng of the indicated GagPol, 300ng CNCG (or CSGW with HIV GagPol), 100ng VSV envelope, and variable amounts of APOBEC as indicated. Empty pCR3.1 was used to normalize the total amount of DNA in each sample. The resulting virions were used to infect MDTF cells overnight with 5μg/mL polybrene and, as before, infection was evaluated according to GFP expression using the Guava Easycyte. Lysates from virus-producing cells were ran onto 4-12% NuPAGE Bis-Tris gels (Invitrogen) and protein was transferred to nitrocellulose membranes followed by probing with an antibody against MLV capsid (262).

HIV evolution in vitro

The infectious titer of both initial viral stocks and supernatent from infected T-cell lines was measured using the same general protocol. 10^4 MT2 or CEMSS cells were seeded in 100μL in each well on a 96-well plate. Five, five-fold serial dilutions were prepared and 50μL of each were added to cells. Following an overnight infection, further rounds of replication were prevented using 100μg/mL dextran sulfate, a cell-entry inhibitor (263). At two days post-infection (dpi), cells were fixed in 2% paraformaldehyde and GFP-positive cells were quantified by FACs.

Prior to selection, HIV-1 Vif+ or ΔVif (described above) were grown in MT2 or CEMSS cells. In 25cm^2 tissue culture flasks, 5x10^5 cells were infected with 500IU in 6mL of total volume and left overnight before a two-fold dilution with fresh media. At 4 dpi (days post infection) 10mL fresh media was added and at 6 dpi cells were transferred to 75cm^2 flasks with 40mL fresh media. Cells were split 1:3 every three days. To maintain the MT2/Vif+ culture during virus-induced cell death, uninfected MT2 cells were added at follows: 5x10^5 cells at 4 dpi, 10^6 cells at 6 dpi and 8 dpi, 3x10^6 cells at 10 dpi, and 4x10^6 cells at 13 dpi. In the case of both CEMSS infected cultures, 10^6 CEMSS cells were added at 8 dpi and 10 dpi and then an additional 4x10^6 cells at 12 dpi. 100μL aliquots were fixed periodically in 2% paraformaldehyde and stored at 4°C. Fifteen days after infection, 50mL of each culture were spun on a tabletop centrifuge at 3000rpm for 15 minutes. The resulting supernatent was filtered through 0.22μm filters and 1.25mL aliquots were frozen at -80°C for further analysis. Cells were washed in PBS and their genomic DNA was extracted for PCR using the Qiagen DNeasy kit.

To amplify sequences packaged into virions, viral RNA was extracted from pelleted virus. In Beckman 14x89mm polyallomer centrifuge tubes, 6mL of filtered T cell supernatent was gently layered on top of 4mL of PBS with 20% sucrose. Samples were spun at 27000rpm in a Beckman ultracentrifuge for two hours at 4°C. The supernatent was aspirated to ~250μL and this remaining volume was shook from the tubes before allowing them to drain upside down for several minutes. The walls of the tubes were dried with kimwipes. Virions were resuspended in 100μL PBS by first incubating on ice for 90 minutes and then pipetting up and down 50 times. RNA was extracted using the Qiagen viral RNA mini kit and cDNA was prepared using the Invitrogen Superscript III kit and oligo-dT primers. Both this cDNA and the genomic DNA from infected CEMSS and MT2 cells were used as PCR template to amplify the polymerase domain of HIV-1 reverse transcriptase (forward primer: 5'-CACTTTAAATTTTCCMATTAGTCCTATT -3' reverse primer: 5'-TGGCAGCACTATAGGCTGTACTG -3'). The resulting products were gel purified and cloned using the Invitrogen TOPO zero-blunt cloning kit. Plasmid clones were sequenced using the Genewiz SP6 primer.

AZT, nevirapine, nelfinavir, and 3TC (provided by the NIH reagent program) were all dissolved in DMSO to 100mM. Drugs were titrated in T cell lines by first infecting four sets of 10^6 CEMSS or MT2 cells with HIV-1 Vif+ at MOI 0.001 in 5mL (eight infections total). After an overnight incubation, cells were split into 12 wells on 24-well plates in media with or without varying concentrations of each drug in 1mL per well. Cells were collected at various time points by fixing 100μL in 2% PFA for later quantitation of GFP+ cells. Cells were split 1:5 every four days with each drug

concentration maintained. Selection using patient plasma was done in the same fashion as the drug titrations, in the 24-well format to conserve plasma. The patient blood sample was collected in heparin-coated tubes. After inactivation at 57°C for 30 minutes, 30μL aliquots of blood plasma were prepared to avoid any additional freeze-thaw cycles after storage at -80°C.

Selections in MT2 cells using AZT and/or nevirapine were done by infecting $4x10^6$ cells at 0.001 MOI in $25cm^2$ tissue culture flasks with supernatants (Vif+ or ΔVif) collected at 15 dpi from MT2 cells (described above). One day after infection, cells were split into plain media or media containing 4μM AZT, 4μM nevirapine, or a combination of 4μM AZT and 500nM nevirapine. Cultures were split 1:5 every four days and enough fresh drug was added to maintain a constant drug concentration with the conservative assumption that no active drug still remained in the culture. This was done to prevent a gradual decrease in drug concentration as degradation occurred over long periods of time. The MT2 culture infected with Vif+ and lacking drug was maintained by the addition of cells every three to four days (or when significant cell death was apparent). Selection using nevirapine in CEMSS was carried out in the same fashion except with a different initial procedure. $2x10^6$ CEMSS cells were infected at 0.002 MOI with supernatent collected from MT2 or CEMSS cells infected 15 days prior with Vif+ or ΔVif virus. At 3 dpi, 500nM nevirapine was added to half of the infected cells.

Receptor Screen

Twenty pools of $2x10^5$ PGSA were transduced with the HeLa cDNA library (MOI 0.5) one day after plating in 6-well dishes. Cells were expanded to 15cm dishes the

following day and selection with 200µg/mL zeocin began that evening. Each pool was grown under selection for six days before plating in duplicate at 10^5/well on a 12-well dish. The following day, cells were infected with CERV2 enveloped virions carrying a CMV-Neo MLV vector (MOI 3.5). One uninfected and two minus-library controls were carried forward in parallel. Each of the 43 groups of cells were expanded to 100mm dishes one day after infection and one day prior to the addition of 1mg/mL G418. Following ten days of G418 selection, colonies were pooled as 20 separate groups corresponding to the original 20-pool PGSA cell expression library. $2.5x10^4$ cells from each of these colony pools were infected independently with CERV2 enveloped LHCX virions in a 48-well plate (MOI 1.0). One day after infection, cells were transferred to 6-well plates and then placed under selection with 100µg/mL hygromycin on the following day. Genomic DNA was extracted from surviving cells using the Qiagen DNeasy kit and used as PCR template with primers designed to anneal to DNA flanking the LMN8iresZeo cloning site (forward primer: 5'- ATACACGCCGCCCACGTGAAGG -3' reverse primer 5'- GCTTCCTTCACGACATTCAACAGACC -3'). The resulting products were cloned using the Invitrogen TOPO zero-blunt cloning kit and four colonies from each were sequenced using the SP6 primer.

Virion Binding

GFP-labeled virus particles were generated by overnight PEI co-transfection of 293T cells with 5µg MLV Gag-GFP expression plasmid (a gift from David Perez-Caballero) and 5µg PCAGGS-CERV2envelope or 5µg PCAGGS on 100mm dishes. The virion-containing supernatant was harvested two days post-transfection and purified

through a 0.22μm filter. CHO cells or a single-cell clone of CHO-humanCTR1 were seeded two days prior to binding onto a poly-lysine coated 15mm coverslip within a 35mm dish at 10^5 cells/dish. Cells to be used for virion binding in the presence of copper were incubated with 20μM $CuCl_2$ for two hours at 37°C. All cells were then incubated in media containing 20mM HEPES pH 7.2 at 4°C for 20 minutes. 980μL virion-containing supernatent was supplemented with 20mM HEPES pH 7.2, 5mg/mL polybrene, and 20μM $CuCl_2$ or water in 1mL of total volume, applied to aspirated cells at 4°C, and incubated at 4°C for one hour. The cells were washed three times with filtered PBS at 4°C before fixation for 20 minutes at room temperature with 4% paraformaldehyde. The cells were then washed twice with PBS.

RNA-interference

A Dharmacon siGENOME smart pool was used to knockdown CTR-1 and included RNA designed to target the following sequences within the CTR1 open reading frame: 5'-AGGCAGUGGUAGUGGAUAU-3', 5'-CUGCGUAAGUCACAAGUCA-3', 5'-GGAACCAUCCUUAUGGagA-3', and 5'-CAUAUGGGGAUGagCUAUA-3'. Cells were seeded on 24-well dishes 24 hours prior to transfection at 10^5cells/well. Media was replaced with Gibco Optimem 2 hours prior to transfections and then again 20 minutes before transfections. Invitrogen Lipofectamine 2000 was used to co-transfect cells with 10pmol of pooled siRNA and 300ng of DsRED expression plasmid (transfection efficiency marker) according to the manufacturer's instructions. Following 6 hours of transfection, cells were supplemented with 1mL DMEM 10% FBS without antibiotics. The transfection was repeated the following day. One hour after the addition of fresh

media to the second transfection, cells were trypsinized and seeded on a 96-well plate at 10^4cells/well for infection the following day. To validate RNAi activity, transfections were repeated as described above except 500ng of PCR3.1-HA-CTR1 was included in the first of two siRNA transfections. Cells were collected for western blot in SDS-PAGE loading buffer two days after the first transfection.

Chapter III. Were TRIM5 and APOBEC3 proteins capable of restricting CERV1 and CERV2?

CERV1 and CERV2 capsid N-terminal domains

To investigate the host range limitation and extinction of chimpanzee endogenous retrovirus groups one and two (CERV1 and CERV2), it was first determined whether TRIM5 restriction factors could have been capable of targeting them during their exogenous replication. Toward this purpose, we studied CERV1 and CERV2 capsid N-terminal domain (CA-NTD) sequences found in the chimpanzee genome as well as their homologues in the rhesus macaque (RhERV1 and RhERV2). The CA-NTD appears to contain all determinants of TRIM5α recognition; this is intuitive given that its highly conserved structure comprises the surface of the retroviral core, which is exposed to TRIM5α following entry into a target cell (108-111, 113, 114). Therefore, CERV1 and CERV2 CA-NTDs in the context of functional chimeric virions are sufficient to test whether TRIM5α was able to restrict these viruses. An ability to target CERV1 and CERV2 capsids would indicate that TRIM5α proteins could have been capable of restricting transmissions during their exogenous replication. The use of a minimal portion of the endogenous retroviruses also increased the likelihood of including intact and functional open reading frames from proviruses that contained frequent mutations. Endogenous murine leukemia viruses (enMLV) were also included for comparison.

All sequences were compiled by BLAST searches of the *Pan Troglodytes, Macaca mulatta* and *Mus musculus* genome sequence databases and were then used to

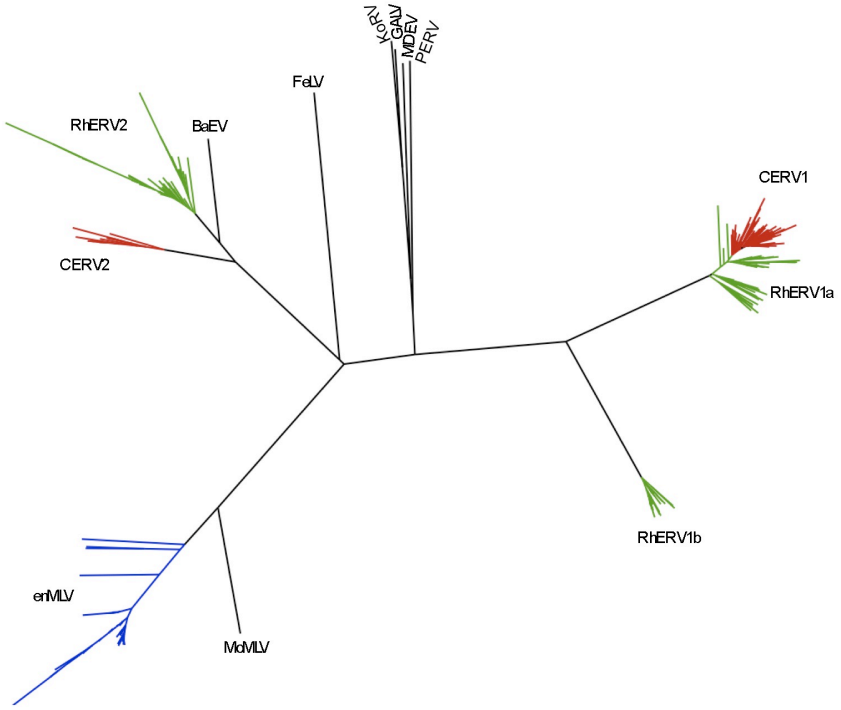


Figure 6. Phylogenetic tree of ancient primate gammaretroviral capsid NTDs. A sequence alignment, and phylogeny of all capsid NTDs from CERV1, CERV2, RhERV1a, RhERV1b, RhERV2, and enMLV and various prototype gammaretroviruses was generated using ClustalX software. The tree diagram was generated and edited using FigTree software.

derive group-specific capsid consensus sequences. A phylogeny of all retrieved endogenous CA-NTDs, as well as those from several other gammaretroviruses, was constructed (Figure 6). A monophyletic group of 85 unique CERV1 CA-NTD sequences was found in the chimpanzee genome, which were closely related to 48 sequences found in the rhesus macaque (RhERV1a). A second, more distant group of 28 CERV1 homologues was also compiled from the rhesus macaque genome database and termed RhERV1b. CERV2, which is more closely related to moloney MLV than CERV1, includes ten CA-NTD sequences in the chimpanzee genome and 44 orthologues in the rhesus macaque (RhERV2). CERV2 and RhERV2, based on their capsid sequences, appear to be closely related to the baboon endogenous retrovirus (BaEV). The mouse genome database contained 41 unique MLV CA-NTD sequences that are from three

groups of endogenous retroviruses (enMLV) previously described as xenotropic, polytropic, and modified polytropic MLV (72).

MLV Chimeric Viruses containing primate and mouse endogenous gammaretroviral CA-NTDs are generally resistant to TRIM5α proteins

To test ancient retroviruses for sensitivity to TRIM5 proteins, functional virions were constructed that contained CA-NTDs from endogenous retroviruses. Libraries of chimeric MLV GagPol were constructed by amplifying CA-NTDs from chimpanzee, rhesus macaque, and mouse genomic DNA and cloning the resulting products in place of the native MLV CA-NTD. It is possible that diversity in such a library is not only the result of natural sequence variations, but could also contain contributions from recombination of highly similar sequences or DNA polymerase errors during PCR. Therefore, in order to be conservative, only sequences that perfectly matched those found in the sequence database were included in the experiments below.

Mus dunni tail fibroblasts (MDTF), that were either unmanipulated or made to stably express TRIM5α proteins, were infected with MLV virions carrying recombinant GagPol proteins. MDTF were chosen due to the lack of a TRIM5α gene in mice. The CA-NTD of the TRIM5α resistant moloney MLV GagPol was replaced with that from N- or B-tropic MLV, or with endogenous retroviral CA-NTDs. The resulting recombinant MLV particles also carried a minimal MLV genome encoding a GFP reporter and were enveloped with VSV-G to allow efficient infection of MDTF. This system allowed for only a single cycle of replication and infections were quantified based on the percentage of MDTF expressing GFP.

As expected, MLV particles carrying the CA-NTD from an N-tropic, but not B-tropic, MLV were restricted by the human, chimpanzee, African green monkey TRIM5α proteins and also, to a lesser degree, by the rhesus macaque TRIM5α. These restricted infections served as controls to demonstrate the functionality of the stably expressed TRIM5α proteins, as well as to confirm that swapping the CA-NTD was sufficient to confer TRIM5α sensitivity. Virions containing CA-NTDs that matched the majority consensus sequence of CERV1 and CERV2 were both infectious (~10^4 IU/mL) and were not restricted to any degree by human, chimpanzee, rhesus macaque, or African green monkey TRIM5α, or by owl monkey TRIM-cyclophilinA (Figure 7A). It should be noted that these results are in sharp contrast to those reported by Kaiser *et al.*, in which both an ancestral CERV1 full-length capsid (both NTD and CTD) and p12 were included in an MLV chimera and this virus was restricted by human and chimpanzee TRIM5α. Aside from including a larger portion of CERV1, the ancestral CERV1 capsid used by Kaiser *et al.* differed by one amino acid from the CERV1 consensus CA-NTD used here (glutamine instead of arginine at position 35). To address this discrepancy, recombinant MLV GagPol genes were constructed that contained full-length CERV1 capsids with either the ancestral or majority consensus sequences. Importantly, inclusion of the entire CERV1 CA did not confer any TRIM5α sensitivity to the chimeric MLV virions (Figure 7B). Furthermore, TRIM5α resistance was observed regardless of whether the consensus or the ancestral CERV1 capsid was used (glutamine or arginine at residue 35, respectively).

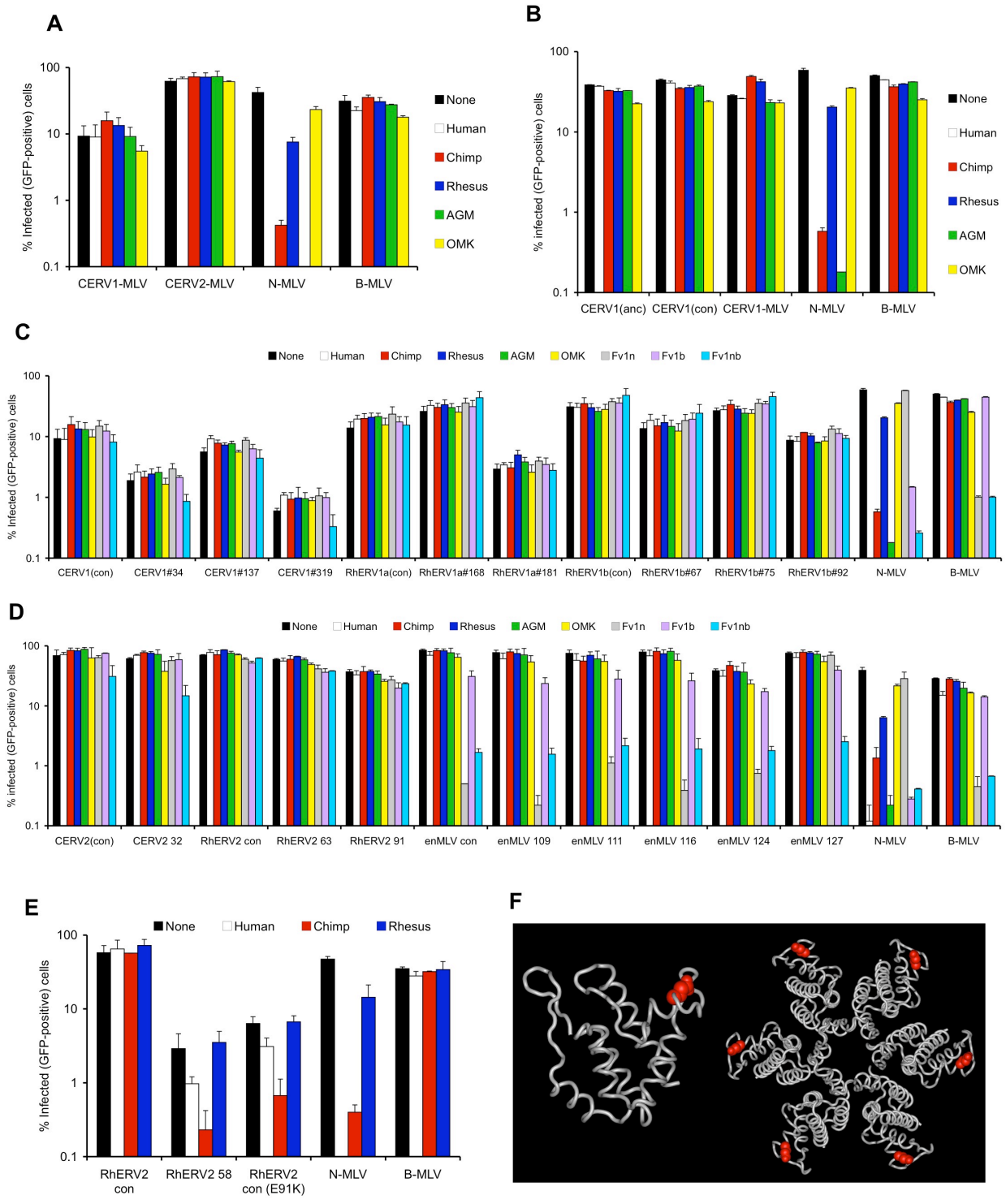
MLV recombinants containing consensus CA-NTDs as well as several functional variants from CERV1, CERV2, RhERV1a, RhERV1b, RhERV2 and enMLV were tested

for sensitivity to TRIM5 proteins as well as to the murine capsid-dependent restriction factors, $Fv1^N$ and $Fv1^B$ (Figure 7C, D). As was described above for particles containing consensus CERV1 and CERV2 CA-NTDs, nearly all chimeric MLV pseudotypes tested displayed resistance to restriction by a variety of primate TRIM5α proteins and owl monkey TRIM-Cyp. In addition, all primate endogenous CA-NTDs were also insensitive to Fv1. With the exception of one variant, the enMLV CA-NTDs conferred a B-tropism to recombinant particles in that they were sensitive to $Fv1^N$ and $Fv1^{NB}$, but insensitive to $Fv1^B$ and TRIM5 proteins.

Only one ancient CA-NTD displayed a TRIM5-sensitive phenotype. This exception to the general TRIM5 resistance observed in all other recombinants was an MLV/RhERV2 recombinant (#58) that was restricted by chimpanzee TRIM5α and marginally by human TRIM5α, but was insensitive to rhesus macaque TRIM5α (Figure 7E). This particular RhERV2 CA-NTD contained four amino acid differences relative to the RhERV2 consensus; one of which, E91K, was not carried by any of the other RhERV2 CA-NTDs tested. On the MLV CA-NTD structure, residue 91 is predicted to lie on the exterior face of the mature core, making it a reasonable putative determinant of TRIM5α sensitivity (Figure 7F). Indeed, when the E91K mutation was introduced into the RhERV2 consensus CA-NTD, the resulting MLV recombinant displayed the same TRIM5α sensitivity as variant #58 (Figure 7E). Therefore, residue 91 is a novel determinant of TRIM5α restriction. Interestingly, the E91K mutation is shared by ~20% of all RhERV2 capsids reported in the genome database, but it is not found in CERV2. This, coupled with the observation that the E91K mutant displayed sensitivity to chimpanzee but not rhesus macaque TRIM5α, suggests that a selective pressure may

Figure 7. Resistance of chimeric MLVs containing primate gammaretroviral CA-NTDs to primate TRIM5 proteins A) MLV chimeric viruses encoding CA-NTD consensus sequence from CERV1 and CERV2 were used to infect MDTF cells expressing human, chimpanzee, rhesus or African green monkey TRIM5α or Owl Monkey TRIMCyp proteins. GFP positive cells were quantified by FACS (Guava EasyCyte) 48h after infection. N-MLV and B-MLV were used as controls. B) MLV chimeric viruses containing an entire ancestral CERV1 capsid sequence (CERV-1(anc)), an entire consensus CERV1 capsid sequence (CERV1(con)) or a consensus CERV1 CA-NTD fused to an MLV CTD (as in A) were used to infect MDTF cells expressing the indicated TRIM5 proteins. N-MLV and B-MLV were used as controls and infection was evaluated as in A). C) MLV-chimeric viruses encoding natural variant CA-NTDs from CERV1, RhERV1a, or RhERV1b were used to infect MDTF cells expressing various TRIM5 and Fv1 proteins. D) MLV-chimeric viruses encoding natural variant CA-NTDs from CERV2, RhERV2 and enMLV were tested as in A). E) MLV chimeric viruses encoding a consensus RhERV2 CA-NTD, a natural variant CA-NTD that includes the E91K mutation (RhERV2 58), or an E91K point mutant were tested as in A). F) Representation of the MLV CA-NTD with the positional equivalent of the RhERV2 E91 residue colored red. The left picture shows a side view of the CA-NTD monomer (outer surface of the capsid at the top) while the right picture shows a view of a single CA-NTD hexamer, viewed from the outside of what would be a complete core.

Figure 7


A
% Infected (GFP-positive) cells
100
10
1
0.1
CERV1-MLV
CERV2-MLV
N-MLV
B-MLV
None
Human
Chimp
Rhesus
AGM
OMK
B
% infected (GFP-positive) cells
CERV1(anc)
CERV1(con)
CERV1-MLV
N-MLV
B-MLV
C
None
Human
Chimp
Rhesus
AGM
OMK
Fv1n
Fv1b
Fv1nb
% Infected (GFP-positive) cells
CERV1(con)
CERV1#34
CERV1#137
CERV1#319
RhERV1a(con)
RhERV1a#168
RhERV1a#181
RhERV1b(con)
RhERV1b#67
RhERV1b#75
RhERV1b#92
N-MLV
B-MLV
D
% infected (GFP-positive) cells
CERV2(con)
CERV2 32
RhERV2 con
RhERV2 63
RhERV2 91
enMLV con
enMLV 109
enMLV 111
enMLV 116
enMLV 124
enMLV 127
N-MLV
B-MLV
E
% Infected (GFP-positive) cells
RhERV2 con
RhERV2 58
RhERV2 con (E91K)
N-MLV
B-MLV
F

have existed in the chimpanzee lineage that was able to prevent the transmission and/or emergence of this particular mutation. In general, though, the TRIM5α resistance displayed by all the other CA-NTD variants tested suggests that it is unlikely that TRIM5α was able to impose complete transmission barriers against the CERV1 and CERV2 families. In particular, the human TRIM5α displayed only marginal activity against RhERV2 capsids carrying the E91K mutation and was completely inactive against all others, which argues against a general protective role for human TRIM5α during potential transmissions of CERV1 and CERV2 to the human lineage.

Evidence of endogenous gammaretrovirus restriction by APOBEC3 family members

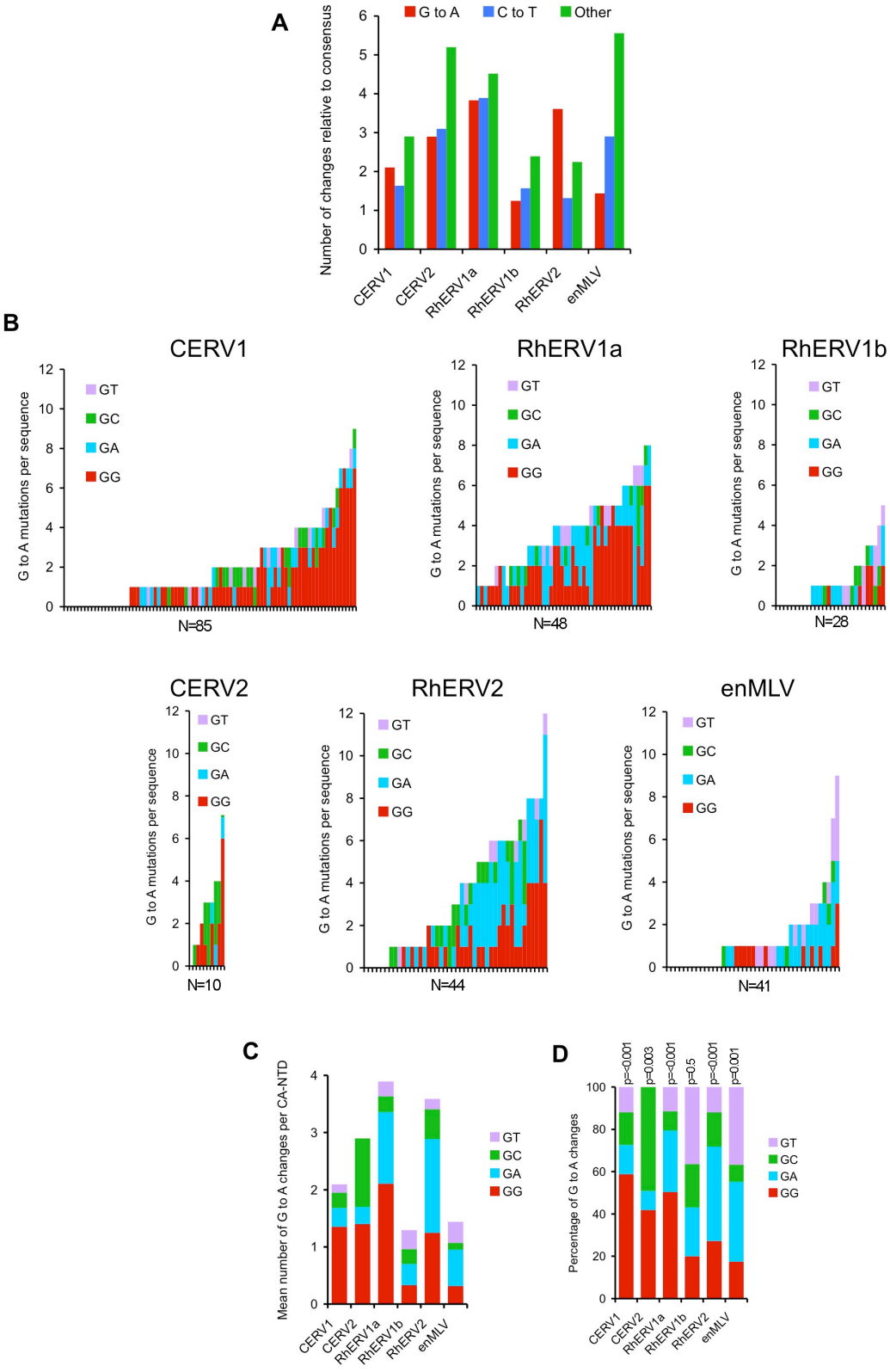
Several APOBEC3 cytidine deaminases are capable of restricting retrovirus replication and a hallmark of their activity is the production of G to A hypermutation within specific sequence contexts. A Bias in CERV1 and CERV2 sequence variability toward these APOBEC3-associated mutations would suggest that these viruses were subject to just such an activity during their exogenous replication. This would suggest a potential role for APOBEC3 restriction factors in the extinction and/or inter-species transmissions of exogenous CERV1 and CERV2. In order to search for signs of past APOBEC3 activity, majority consensus sequences were derived for CA-NTDs from CERV1, RhERV1a, RhERV1b, CERV2, RhERV2, and enMLV. Within each of these groups, any deviation from the consensus sequences was documented. A striking degree of conservation was observed in all groups of CA-NTDs with a typical divergence from a

given majority consensus sequence of <3%. In several groups, though, there existed a high relative abundance of G to A and C to T mutations (Figure 8A). Furthermore, the relative levels of G to A and C to T mutations varied both between homologues and within a given host species. In RhERV2 sequences, half of all nucleotide differences found in the CA-NTDs were G to A mutations. On the other hand, G to A mutations in CERV1, CERV2, RhERV1a, and RhERV1b CA-NTDs occurred at frequencies that were similar to C to T mutations and both were present at a lower level than all other mutations pooled together. Also in contrast to the mutations in RhERV2 CA-NTDs were those found in enMLV CA-NTDs, in which the majority were from neither G to A nor C to T substitutions.

The most common nucleotide mutations found in mammalian genomes are generally G to A and C to T transitions, which are typically caused by spontaneous deamination (described below). APOBEC3-induced deamination of viral reverse transcripts can be distinguished from spontaneous deamination of the host genome by the existence of certain biases in the dinucleotide context of the mutations. Namely, APOBEC3G specifically deaminates CC dinucleotides on the cDNA minus strand to CU, whereas primate APOBEC3F, 3B, and 3H and murine APOBEC3 all mutate TC to TU (153-155, 160, 170, 171)**.** These changes, made during reverse transcription, manifest in the positive strand sequence as GG to AG or GA to AA mutations. Additionally, rare deamination of cytidine residues on HIV-1 RNA by APOBEC1 has been reported (194).

Figure 8. Analysis of ancient ERV CA-NTDs. A) The mean number of G to A, C to T, and all other changes (pooled) per CA NTD sequence, relative to each consensus sequence is shown. B) The numbers of G to A mutations for each individual CA-NTD recovered from sequence databases is plotted as a bar graph (one bar for each provirus) and color-coded according to dinucleotide context (i.e. the nucleotide in the +1 position relative to the G to A change; red = GG to AG, cyan = GA to AA, green = GC to AC, magenta = GT to AT). The CA NTD sequences are arranged from left to right in order of increasing numbers of total G to A changes. C) The mean number of G to A changes per CA-NTD sequence, relative to the consensus, is plotted for pooled data for each virus species and color coded according to dinucleotide context as in B. C) The percentage of G to A changes in each dinucleotide context relative to consensus sequence is plotted, and normalized according to the dinucleotide composition of the species-specific consensus sequences. The p-values for deviation from a random distribution were calculated using the chi-squared test.

Figure 8


A
G to A
C to T
Other
Number of changes relative to consensus
CERV1
CERV2
RhERV1a
RhERV1b
RhERV2
enMLV
B
CERV1
RhERV1a
RhERV1b
CERV2
RhERV2
enMLV
GT
GC
GA
GG
G to A mutations per sequence
N=85
N=48
N=28
N=10
N=44
N=41
C
Mean number of G to A changes per CA-NTD
D
Percentage of G to A changes
p=<0.001
p=0.003
p=<0.001
p=0.5
p=<0.001
p=0.001

In order to ask whether the excessive G to A and C to T changes found in the endogenous gammaretroviruses were caused by spontaneous deamination of proviruses in the host genome or by the catalytic action of antiviral deaminases, the dinucleotide context of these mutations was quantified to detect any biases that may have existed. Because the number of mutations observed in each individual CA-NTD was rather small, we inspected both the pooled CA-NTD sequence data from each viral species (Figure 8C) as well as the individual CA-NTD sequences (Figure 8B). There was significant variation in the burden of G to A mutations among individual CA-NTDs: while some completely lacked G to A mutations, others contained up to 12 G to A mutations in the ~400 nucleotide CA-NTD encoding sequence (Figure 8B). Moreover, striking and variable biases in the patterns of G to A mutations were observed. For example, in CERV1 CA-NTDs, G to A changes occurred primarily in the context of GG dinucleotides (Figure 8B, C). This bias was statistically significant, and was maintained when corrected for the frequency with which each GN dinucleotide occurred in the consensus sequence (Figure 8D). In CERV2 CA-NTDs, the occurrence of G to A changes was biased toward both GG and GC dinucleotides; however, in this case the overall numbers of G to A changes were small (Figure 8B) because only 10 CERV2 proviruses are present in chimpanzee DNA. Interestingly, compared to each of their chimpanzee counterparts, the rhesus macaque endogenous gammaretroviruses RhERV1a and RhERV2 exhibited a different mutational bias: G to A mutations occurred frequently at both GG and GA dinucleotides, and these biases were also highly statistically significant (Figure 8B, C). In contrast, RhERV1b CA NTDs were quite different in that they had relatively few G to A mutations and their occurrence was unbiased with respect to dinucleotide context (Figure 8B).

Confirmation of mutational biases in endogenous gammaretroviruses

The sequence analyses above suggested that some endogenous proviruses may well have encountered antiviral cytidine deaminases. Therefore, to confirm and extend these observations, we asked if similar patterns of mutation could be observed in another region of the endogenous proviruses, in particular envelope (Env) sequences (Figure 9). Because there are large numbers of CERV-1 proviruses and the chimpanzee genome database contained abundant unresolved regions, we retrieved only envelope sequences that were linked to CA-NTDs that had either ≤1 G to A mutations or ≥4 G to A mutations (10 of each). Overall, there was a good correlation between the CA-NTD and Env sequences in terms of the burden of the G to A mutations and the dinucleotide context in which they occurred (Figure 10 A, B). Specifically, the occurrence of GG to AG mutations in a given CERV1 CA-NTD strongly predicted the occurrence of the same type of mutation in the linked Env sequence, while CA-NTD sequences that contained one or no G to A mutations were linked to envelopes that contained relatively few G to A mutations that largely lacked a dinucleotide bias. Indeed, there was only one exception to this finding among the 20 CERV-1 sequences analyzed, in which a CA-NTD bearing three GG to AG mutations was linked to an Env sequence containing a light and unbiased burden of G to A changes (Figure 10 A, B).

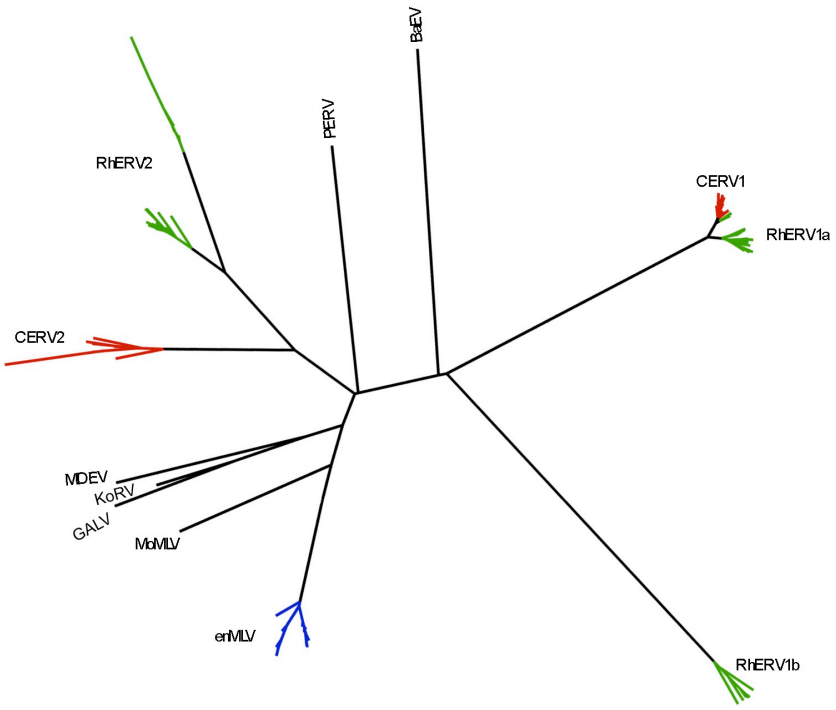


Figure 9. Phylogenetic tree illustrating the relationship between endogenous primate gammaretroviral Env sequences and other gammaretroviral Envs. A sequence alignment, and phylogeny of Env sequences from CERV1, CERV2, RhERV1a, RhERV1b, RhERV2, and enMLV and various prototype gammaretroviruses was generated using ClustalX software. The tree diagram was generated and edited using FigTree software.

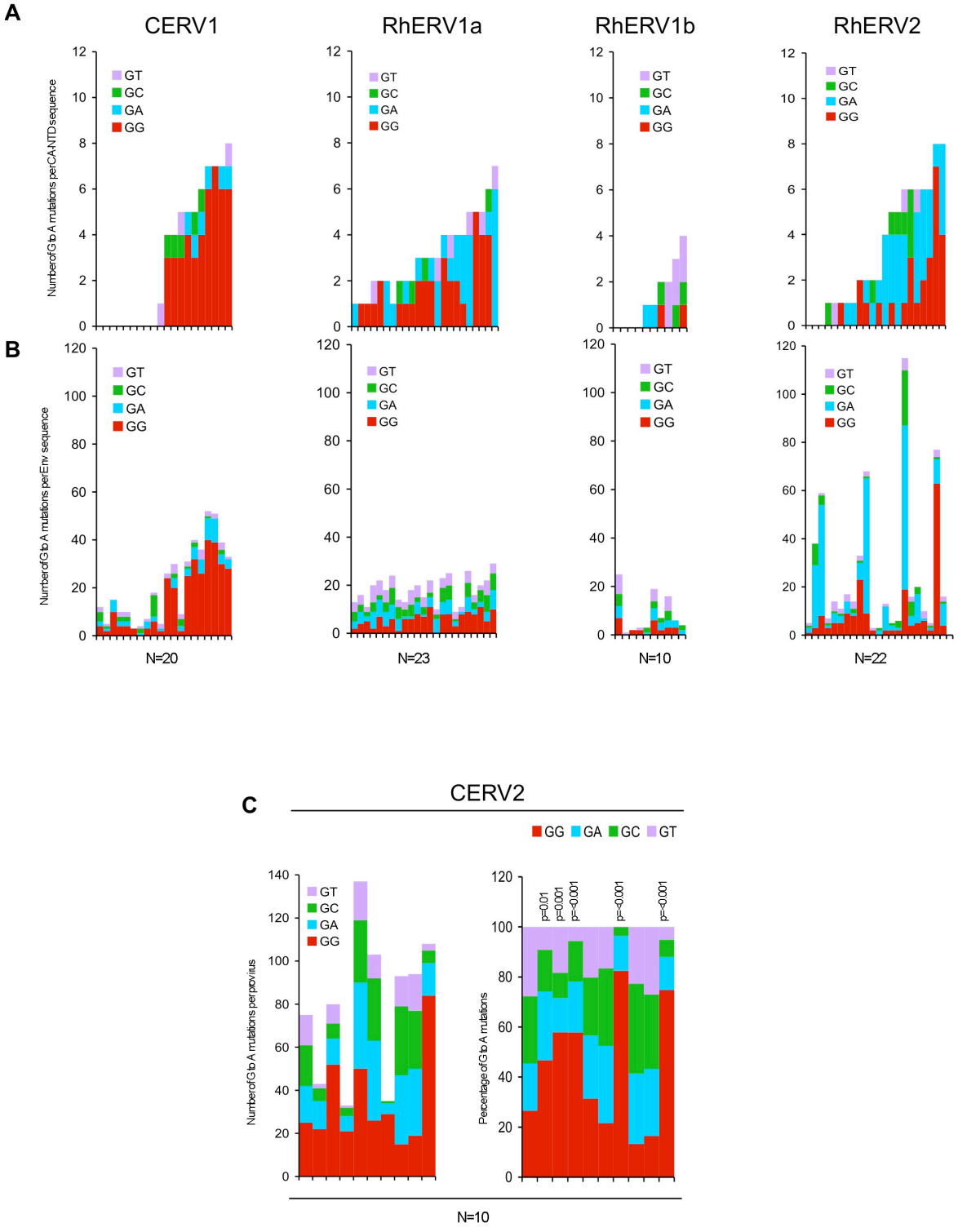
Because there are only ten CERV2 proviruses in the chimpanzee genome (and only four of them are complete, with others ranging in size from 2.6 to 7.7 kb) we took a different approach to confirm or refute the notion that they contained excessive or biased G to A mutations. Specifically, we analyzed all of the available sequence for all ten proviruses. In part because the CERV2 proviruses were intrinsically divergent, and differed in length, the absolute numbers of G to A changes as compared to the consensus sequence was quite variable among them (Figure 10C). However, when the character of the G to A mutations was examined, five of the 10 proviruses exhibited a clear,

statistically significant, excess of GG to AG mutations, as compared to overall G to A changes (Fig 10C). Thus, the patterns of nucleotide substitutions strongly suggested that both groups of endogenous gammaretroviruses that integrated into the chimpanzee genome in the past few million years frequently encountered the only mutagen known to preferentially induce GG to AG mutations, namely APOBEC3G, prior to or during endogenization.

The coverage of the rhesus macaque reference genome database was less complete compared to that of the chimpanzee, so only 23, 10 and 22 envelope sequences that were unambiguously linked to CA-NTDs could be retrieved for RhERV1a, RhERV1b and RhERV2 respectively. Moreover, in the case of RhERV2, the CA-NTD linked env genes were distributed between two distinguishable subgroups and, therefore, a separate Env consensus sequence was deduced for each group. For both RhERV1b and RhERV2 the analysis of env sequences corroborated the results obtained using CA-NTDs (Fig 10, B). Specifically, the numbers of G to A changes in RhERV1b Env, like CA-NTD sequences were modest and unbiased with respect to dinucleotide context; while in RhERV2, there was a clear excess of both GG to AG and GA to AA mutations in both CA-NTD and Env sequences (Figure 10A, B). However, while there was concordance with respect to overall context bias of G to A mutations present in RERV2 CA-NTD and Env sequences, inspection of individual linked CA-NTD and Env sequences revealed that the burden and character of G to A mutations was variable. Specifically, several Env sequences bore a large proportion of G to A substitutions in GA dinucleotides while two other Env sequences bore predominantly GG to AG hypermutation. Additionally, the

Figure 10. Analysis of G to A changes in CA-NTD and Env sequences in primate gammaretroviruses. A and B) CA-NTD sequences from the primate gammaretroviruses were selected according to whether an unambiguously linked Env sequence could be retrieved from the genome database. In the case of CERV1, ten sequences that essentially lacked G to A changes and Ten sequences that had four or more G to A changes were selected. For each virus species each sequence is plotted as a bar graph (one bar for each provirus) and color-coded according to dinucleotide context (i.e. the nucleotide in the +1 position relative to each G to A change; red = GG to AG, cyan = GA to AA, green = GC to AC, magenta = GT to AT). The CA NTD sequences are arranged from left to right in order of increasing numbers of total G to A changes in CA-NTD sequences (A). The same analysis was performed for the linked env sequences (B), which are arranged in order according to the linked CA-NTD sequence. C) All ten CERV2 proviral sequences were analyzed and the dinucleotide context in which G to A mutations occurred is plotted and color-coded as in A (left) panel. Additionally, The percentage of G to A changes in each dinucleotide context relative to consensus sequence is plotted, and normalized according to the dinucleotide composition of the consensus sequences (right panel). The p-values for deviation from a random distribution were calculated using the chi-squared test. Proviruses 1, 4, 6 and 7 are complete.

Figure 10


A
CERV1
RhERV1a
RhERV1b
RhERV2
GT
GC
GA
GG
Number of G to A mutations per CA-NTD sequence
B
Number of G to A mutations per Erv sequence
N=20
N=23
N=10
N=22
C
CERV2
GG GA GC GT
Number of G to A mutations per provirus
Percentage of G to A mutations
p=0.01
p=0.001
p=<0.001
p=<0.001
p=<0.001
N=10

degree to which CA-NTDs and Env sequences were mutated in individual proviruses did not always correlate; in some cases hypermutated RhERV2 Env sequences were linked to CA-NTD sequences that bore relatively few G to A changes (Figure 10A, B). In RhERV1a, the nature and burden of G to A changes in Env sequences was not at all predicted by the findings in the CA-NTD sequences. Indeed, the preferential occurrence of G to A changes in Env in the context of GG and GA dinucleotides was marginal, and clearly less pronounced than in the CA-NTD sequences (Fig 10, A, B). These results, and the discordant findings with respect to G to A mutation in a few RhERV2 CA-NTD versus linked Env sequences are either explained by variable mutation frequencies across genomes (a known characteristic of APOBE3G induced mutation) or by recombination between hypermutated and non-hypermutated viral genomes that occurred prior to or during deposition in the germ line. Indeed, inspection of several complete or nearly complete RhERV2 proviruses revealed clear variation in the extent of G to A mutation in some (but not all) proviruses across the viral genome (Figure 11). Notably, an obvious 5' to 3' gradient of increasing mutation intensity was present in some examples bearing excessive G to A mutations, as has previously been reported in hypermutated HIV-1 genomes (171, 172). Also, several recombination break points were implied by abrupt differences in the frequency of APOBEC3-induced mutation.

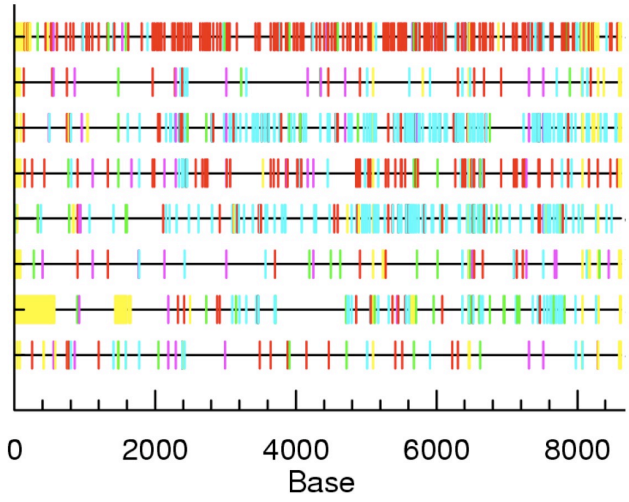


Figure 11. Examples of the RhERV2 proviruses illustrating variation in type, burden, and distribution of G to A mutations among endogenous primate gammaretroviruses. Each horizontal line represents a complete or nearly complete ~8.5kB provirus (nucleotide position scale is given at the bottom of the diagram), vertical marks indicate the position of G to A mutations relative to the RhERV2 consensus sequence and are color coded according to dinucleotide context (i.e. the nucleotide in the +1 position relative to each G to A change; red = GG to AG, cyan = GA to AA, green = GC to AC, magenta = GT to AT). Yellow marks indicated non-G to A mutations, including in/dels.

Biased patterns of G to A mutation in primate ERVs are not explained by spontaneous cytidine deamination

The chimpanzee and rhesus gammaretroviral CA-NTD and Env sequences also contained a striking excess of plus strand C to T changes relative to their respective consensus sequences (Figure 8A). Potentially, the high rate of C to T mutations in CA-NTD sequences could be the result of cytidine deamination after integration of the provirus into the germ line. Correspondingly, excessive plus strand G to A changes could be a reflection of cytidine deamination events on the minus strand within the host genome. In contrast to APOBEC3-mediated deamination events, which occur

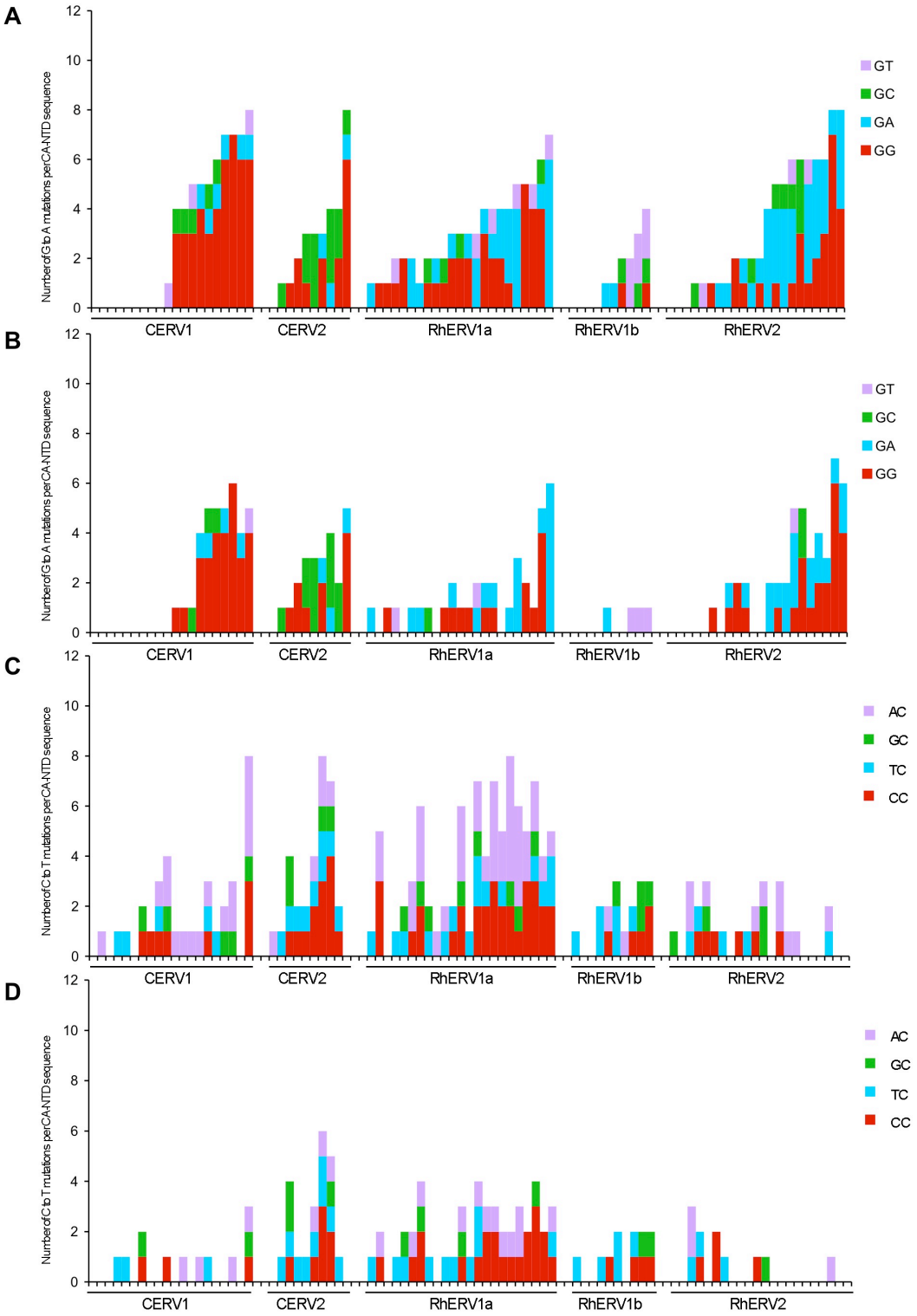
specifically on the anti-sense DNA strand and are profoundly influenced by the identity of the nucleotide in the -1 position relative to the deaminated cytidine, mutations that arise as a result of cytidine deamination in the host genome are influenced by the nucleotide in the +1 position and lack any strand specificity. Specifically, the most common cytidine deamination-induced mutation found in mammalian genomes occurs in the context of CG dinucleotides (i.e. G in the +1 position). This results in CG to TG or CG to CA sense mutations depending on whether CG dinucleotides on the plus or minus strand are deaminated, respectively (264, 265). This phenomenon is explained by a transcriptional control mechanism that operates by cytidine methylation of CG islands. When genomic 5-methyldeoxycytosine is deaminated, deoxythymidine is formed. Therefore, if a CG to TG mutation (or its anti-sense counterpart, CG to CA) is found in an endogenous retrovirus it is likely the result of deamination of proviral nucleotides during long-term residence in the host genome as opposed to deamination by restriction factors during exogenous replication. Interestingly, it is possible that deamination of CG dinucleotides in the cellular genome is catalyzed by the APOBEC family member, activation-induced cytidine deaminase (AID). This is suggested by the observation that AID expression is required for proper demethylation in pluripotent stem cells (266, 267). It was proposed that during reprogramming of gene expression, silenced DNA can be demethylated through the deamination of methyl-cytidine followed by T:G base excision repair by glycosylases. If this mechanism resulted in mutations of endogenous retroviruses, though, it would be due to their presence in the cellular genome rather than the result of a specific antiviral activity.

To account for the potential "noise" introduced by deamination in the cellular genome, we counted and categorized all G to A and C to T changes for both CA-NTD and Env sequences before and after exclusion of CG dinucleotides (Figure 12). Notably, in an analysis of plus strand C to T mutations there were no evident biases in the -1 position after the exclusion of substitutions that occurred in CG dinucleotides (Figure 12C, D). This confirms that the APOBEC3 activities identified had minus-strand specificity. Moreover, exclusion of C to T mutations that occurred in the context of CG dinucleotides on the minus strand, reduced the overall number of plus strand G to A mutations but did not affect conclusions with respect to their associated dinucleotide context biases (Figure 12A, B).

Notably, the relative extent of G to A as compared to C to T mutations varied among individual proviruses of a given virus group as well as between virus groups. This was most evident in the analysis of Env sequences (or complete proviral sequences in the case of CERV2), because the longer length of sequence analyzed permitted a more robust estimate of G to A and C to T mutation frequencies (Figures 13, 14). For many of the individual CERV1, CERV2, and RhERV2 proviruses, the frequency of G to A mutations greatly exceeded the frequency of C to T mutations (Figure 13A, C, Figure 14), and in proviruses where this was the case, the excessive G to A mutation was invariably in the context of GG or GA dinucleotides, and remained evident when CG dinucleotides were purged from the analysis (Figure 13B, D and Figure 14). Conversely, RhERV1a and RhERV1b Env sequences exhibited broadly similar levels of comparatively unbiased G to A and C to T mutation (Figure 13) suggesting that the frequency of APOBEC3-induced mutation was low compared to that from other sources. A comparison with

Figure 12. Comparative analysis of the context in which G to A versus C to T changes occur in primate gammaretroviral CA-NTDs. (A) Each sequence is plotted as a bar graph (one bar for each provirus) and color-coded according to dinucleotide context (i.e. the nucleotide in the +1 position relative to each G to A change; red = GG to AG, Cyan = GA to AA, Green = GC to AC, magenta = GT to AT) the CA NTD sequences are arranged from left to right in order of increasing numbers of total G to A changes in CA-NTD sequences (panel A contains the same data as Figure 10A) (B) Same analysis as in A, except mutations were enumerated after removal of minus strand CG to TG (plus strand CG to CA) mutations. (C) Analysis of plus strand C to T mutations, in the same CA-NTD sequences, in the same order, (left to right) as in A and B. C to T changes are color-coded according to the nucleotide in the -1 position relative to each C to T change; red = CC to CT, Cyan = TC to TT, Green = GC to GT, magenta = AC to AT) (D) Same analysis as in C, except mutations were enumerated after removal of plus strand CG to TG mutations.

Figure 12

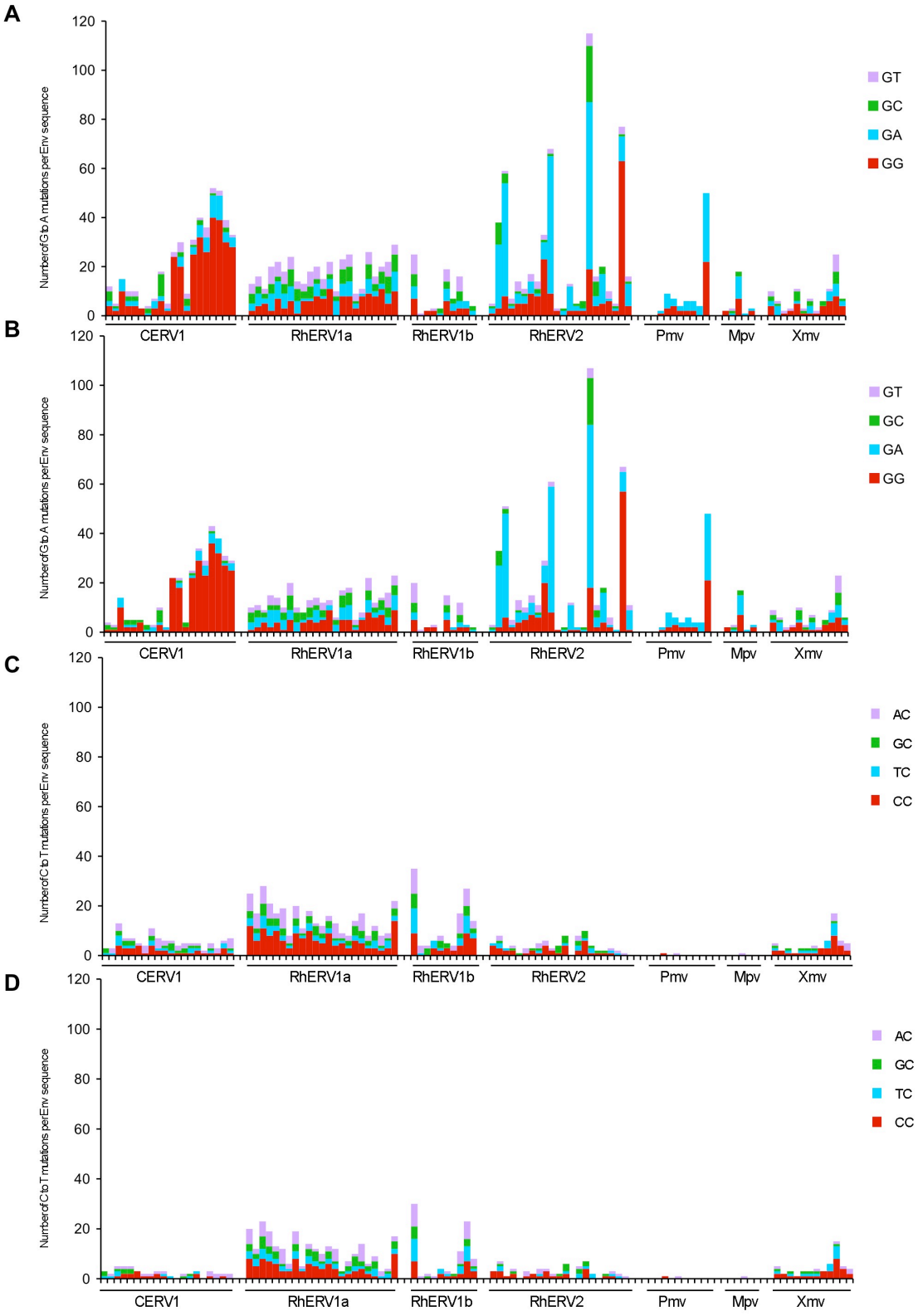

A
B
C
D
Number of G to A mutations per CA-NTD sequence
Number of C to T mutations per CA-NTD sequence
0
2
4
6
8
10
12
CERV1
CERV2
RhERV1a
RhERV1b
RhERV2
GT
GC
GA
GG
AC
GC
TC
CC

endogenous MLVs revealed a similar situation to that in primate gammaretroviruses, and as was previously reported, two out of the three groups of endogenous MLVs (Pmv, Mpv, but not Xmv) exhibited higher levels of G to A than C to T mutations, and this excess in G to A mutations was associated with similar context biases (268). Overall, clear dinucleotide context biases were observed in each situation where G to A mutations clearly outnumbered C to T mutations.

Thus, the excessive and context-biased G to A mutations that are present in some of the groups of primate endogenous gammaretroviruses cannot be explained by spontaneous cytidine deamination. Rather, the analysis of both CA-NTD sequences, and linked envelope sequences, provides robust support for the hypothesis that APOBEC3G in chimpanzees, and a combination of APOBEC3G and other APOBEC3 proteins in rhesus macaques, extensively mutated ancient gammaretroviral genomes prior to and/or during their endogenization. If endogenization involved multiple rounds of replication in reproductive organs, then the high expression levels of APOBEC3 proteins in the ovaries and testes suggest that they may have been able to act on viruses during endogenization (269).

Figure 13. Comparative analysis of the burden of, and the context in which, G to A versus C to T changes occur in endogenous primate and murine gammaretroviral Env sequences. (A) Each sequence is plotted as a bar graph (one bar for each provirus) and color-coded according to dinucleotide context as before. The primate Env sequences are derived from the same proviruses as the CA-NTD sequences shown in Figure 12, and arranged from left to right in the same order (panel A contains the same data as Figure 10B) (B) Same analysis as in A, except mutations were enumerated after removal of minus strand CG to TG (plus strand CG to CA) mutations. (C) Analysis of plus strand C to T mutations, in the same Env sequences, in the same order, (left to right) as in A and B. C to T changes are color-coded according to the nucleotide in the -1 position relative to each C to T change as before. (D) Same analysis as in C, except mutations were enumerated after removal of plus-strand CG to TG mutations.

Figure 13


A
B
C
D
Number of G to A mutations per Env sequence
Number of C to T mutations per Env sequence
0
20
40
60
80
100
120
CERV1
RhERV1a
RhERV1b
RhERV2
Pmv
Mpv
Xmv
GT
GC
GA
GG
AC
GC
TC
CC

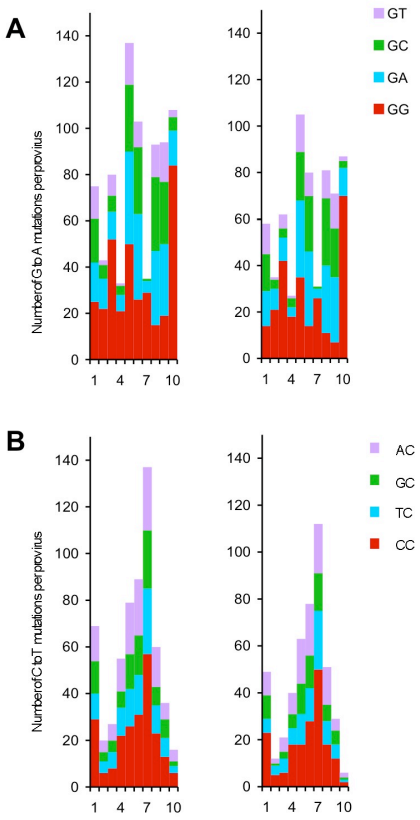


Figure 14. Comparative analysis of the burden of, and the context in which, G to A versus C to T changes occured in CERV2 proviruses. (A) Each sequence is plotted as a bar graph (one bar for each provirus) and color-coded according to the dinucleotide context in which G to A mutations occur, as described above. The proviral sequences correspond to the CA-NTD sequences shown in Figure 12, and arranged from left to right in the same order. The left panel shows analysis without removal of minus strand CG dinucleotides, while the right panel shows analysis after their removal. (B) Analysis of plus strand C to T mutations, in the same CERV2 proviral sequences, in the same order, (left to right). C to T changes are color-coded according to the nucleotide in the -1 position relative to each C to T change as is Figure 12. The left panel shows analysis without removal of plus strand CG dinucleotides, while the right panel shows analysis after their removal.

Effects of cytidine deamination on ancient primate gammaretrovirus CA-NTD function

Inspection of the endogenous primate gammaretroviral CA-NTD sequences revealed that, with the exception of RhERV1b CA-NTD sequences, many harbored stop codons (Figure 15A). Strikingly, G to A mutations generated a large fraction of the stop codons. C to T changes resulted in stop codons somewhat less frequently than G to A changes, but still provided a major source of these obviously inactivating mutations. Frameshifts and other mutations were responsible for only a small proportion of protein truncating mutations. Many of the C to T changes occurred in CG dinucleotides and are therefore likely to be caused by spontaneous post-integration deamination events. Notably, all of the stop codons that arose through the appearance of G to A mutations were due to a change from Trp (TGG) to termination codons (TAG, TGA or TAA), often reflecting the dinucleotide bias associated with the G to A changes observed in those species. The large number of stop codons generated at Trp codons, likely by APOBEC3-mediated cytidine deamination, would almost certainly functionally inactivate many of these ancient endogenous retroviruses. Note that this analysis was confined to the CA-NTD sequences, and thus indicates the predominate source of termination codons, but significantly underestimates the total number of termination codons in full-length proviruses.

In addition to stop codons, G to A and C to T changes introduced missense mutations in CA NTDs whose effects cannot be determined simply by inspecting sequence data. Although it was not practical to determine the effect of each G to A and C

to T mutation in isolation, a few (fourteen) members of the libraries of chimeric CERV1-RhERV1a- or RhERV1b-MLVs bore single, naturally occurring, missense mutations that could likely be attributed to enzymatic cytidine deamination (GG to AG changes) or spontaneous cytidine deamination (CG to CA or CG to TG changes). Chimeric MLVs bearing these single mutations were invariably less infectious than corresponding MLVs encoding the intact CERV1, RhERV1a or RhERV1b consensus CA-NTD sequence (Figure 15B). The magnitudes of the decreases in infectivity were variable, and in one case a missense GG to AG mutation completely inactivated a CERV1-MLV chimera. Thus, amino acid differences from the consensus, likely attributable to enzymatic or spontaneous cytidine deamination decreased infectivity, suggesting that they were unlikely to be positively selected to provide benefit to a replicating virus. Rather, in addition to their propensity to induce the occurrence of termination codons, the abundant nonsynonymous G to A and C to T changes likely contributed significantly to a deleterious genetic load borne by these populations of endogenized viruses.

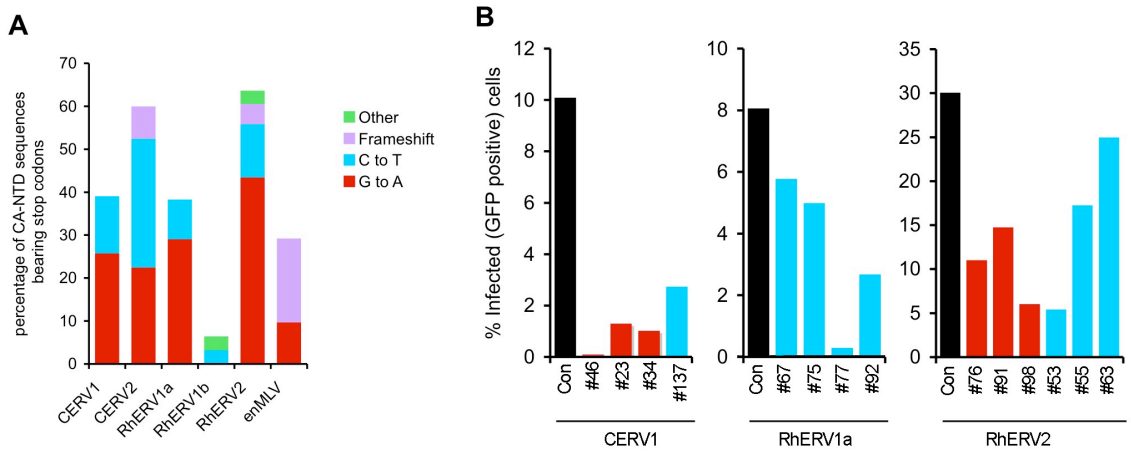


Figure 15. Effects of cytidine deamination on CA-NTD function. A) the percentage of translation termination codons generated by G to A, C to T, frameshift and other mutations in endogenous primate and mouse gammaretroviral CA-NTDs. The percentage of all CA-NTD sequences with stop codons in CERV1, CERV2, RhERV1a, RhERV1b, RhERV2, and enMLV CA-NTDs is plotted and the colored subdivision indicate the proportions that are attributable to each type of mutation. B) Effect of single G to A and C to T missense mutations on the infectivity of chimeric MLVs bearing CA-NTDs from primate endogenous retroviruses. Consensus capsid NTDs from the indicated virus species and variants bearing single naturally occurring G to A mutations (red) or C to T mutations (cyan) were introduced into chimeric MLV GagPol and infectivity measured as in Figure 7. Note that all three G to A mutations in CERV1 (#46, #23, #34) occurred in the context of GG dinucleotides, while the mutations in RhERV1a and RhERV2 were all in the context of CG dinucleotides on the plus strand (blue) or minus strand (red).

CERV2 restriction by APOBEC3 proteins in vitro

It is clear from the analyses presented above that APOBEC3G was capable of heavily mutating CERV2, but this alone does not necessarily indicate that APOBEC3G was able to broadly restrict its replication. Indeed, APOBEC3G is able to induce mutations in HERV-K without restricting a HERV-K consensus during a single cycle of replication (13, 270). Also, in spite of the ability of HIV-1 to avoid restriction by human APOBEC3 proteins through the action of Vif, hypermutated proviruses are abundant in

infected patients (192, 193, 197, 271). Furthermore, HIV-1 can accumulate APOBEC3-associated mutations in cells that allow equal replication of Vif-deficient and Vif-proficient viruses (Chapter IV). These results point out that APOBEC3-induced mutation is observable even when viruses are not susceptible to its restrictive effects. Thus, it becomes necessary to confirm, *in vitro,* that APOBEC3G is capable of restricting CERV2 infection.

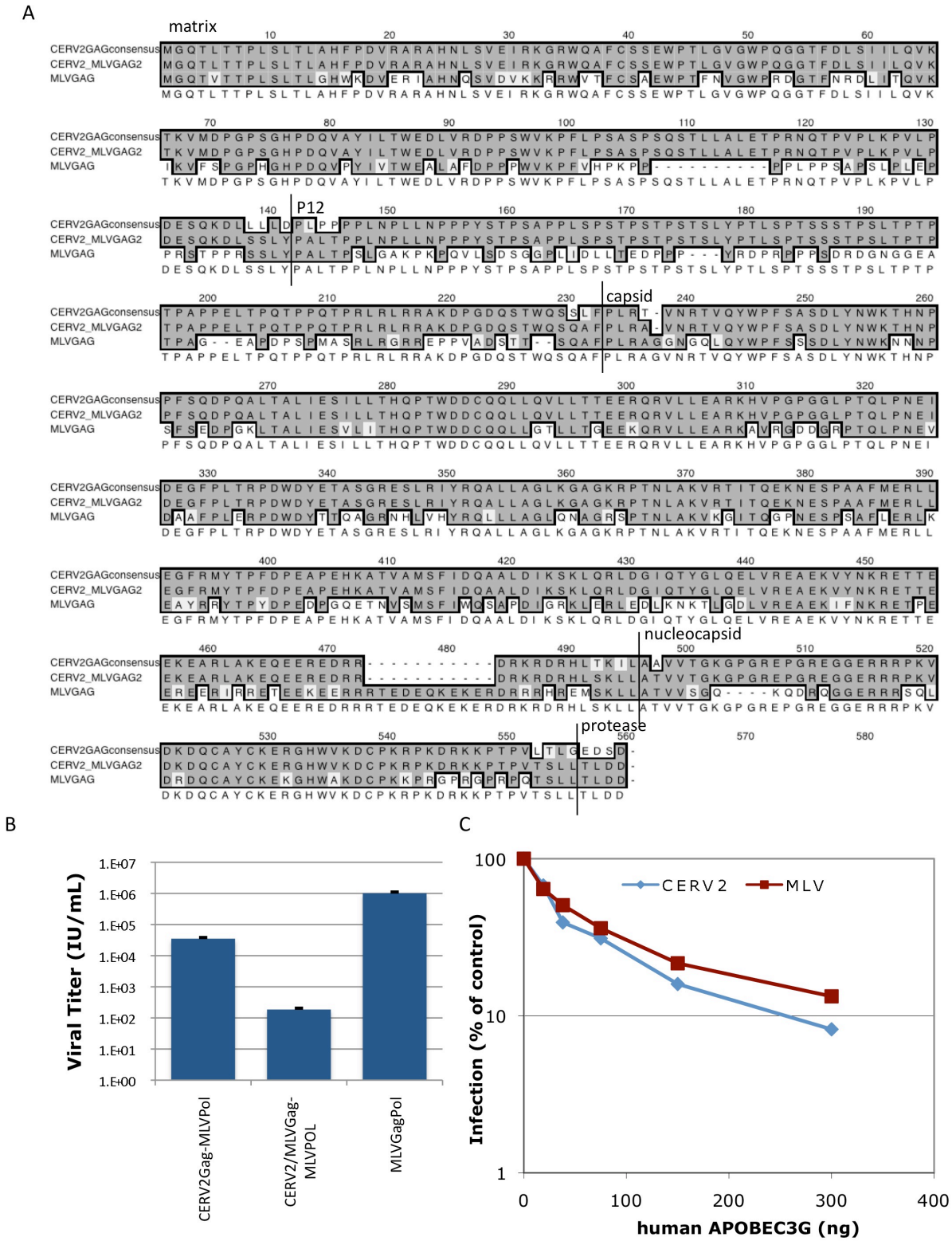
It has been posited that the incorporation of APOBEC3G into HIV particles occurs through its interaction with virion-packaged RNA and nucleocapsid (162-168). Even in the absence of viral RNA, though, APOBEC3G was shown to both incorporate into virions and interact with Gag in an RNA- and nucleocapsid-dependent manner (162, 164, 165, 168). This suggested that APOBEC3 packaging depends, not on specific sequences in the viral genomic RNA, but on nucleocapsid and, in particular, its ability to bind RNA. Further importance of nucleocapsid in APOBEC packaging has been shown for human T cell leukemia virus type 1, which contains a peptide motif at the C-terminus of nucleocapsid that is unique to primate T cell leukemia viruses and is responsible for the exclusion of APOBEC3G from particles (272). Overall, it is clear that the inclusion of the major structural components of CERV2 is warranted in an *in vitro* study of CERV2 restriction by APOBEC3. Thus, we constructed a consensus CERV2 Gag gene and combined it with MLV Pol to make a chimeric GagPol (CERV2Gag-MLVPol). Due to the requirement for MLV protease to process the CERV2 Gag in the maturation of chimeric particles, a second version of CERV2 Gag (CERV2/MLV Gag) was synthesized that replaced four amino residues on both sides of protease cleavage sites with those from MLV (Figure 16A). Both chimeric GagPol proteins were functional in the context of

VSV enveloped virions carrying a GFP reporter and, interestingly, intact CERV2 Gag protease cleavage sites yielded particles with a titer that was two orders of magnitude higher than particles containing cleavage sites from MLV Gag (Figure 16B).

To measure restriction by APOBEC3G, variable amounts of Myc-tagged human APOBEC3G were co-transfected with VSVg, an MLV vector encoding a GFP reporter, and either MLV GagPol or CERV2Gag-MLVPol. APOBEC3G was able to restrict virions containing CERV2 Gag by up to approximately one order of magnitude, which was similar to the restriction of particles carrying MLV Gag (Figure 16C). This result suggests that APOBEC3G was, indeed, capable of restricting CERV2 replication. To address a possible role for APOBEC3G in preventing CERV2 endogenization in the human lineage, it was determined if human APOBEC3G was potentially more potent against CERV2 than orthologues from species whose genomes were invaded by CERV2 family members. VSV-G enveloped CERV2Gag-MLVPol, MLV GagPol, or HIV GagPol virions were generated in the presence of variable amounts of APOBEC3G from human, chimpanzee, or rhesus macaque (Figure 17). Although human APOBEC3G was more potent in terms of its restrictive effects compared to rhesus macaque APOBEC3G, it was weaker than chimpanzee APOBEC3G. Furthermore, this order of APOBEC3G potency was observed in the restriction of all three viruses tested. Therefore, a protection of the human lineage from CERV2 cannot be ascribed to a restriction by APOBEC3G that was potentially more powerful in human ancestors than it was in primates that were unable to avoid CERV2 endogenization.

Figure 16. A functional CERV2 Gag consensus and its sensitivity to human APOBEC3G. A) Alignment of Gag amino acid sequences: CERV2 majority consensus sequence, CERV2 majority consensus sequence with protease cleavage sites (marked with vertical line) matching moloney MLV (CERV2/MLVGag), and moloney MLV. B) VSV enveloped pseudotypes were generated by cotransfection with CNCG and MLV GagPol or chimeras containing CERV2 Gag. The number of infectious units per mL supernatant was measured by infection of 10^4 MDTF with a five-fold serial dilution. C) MLV GagPol or CERV2Gag-MLVPol particles used to infect MDTF were prepared as in (B) except with cotransfection of variable amounts of human APOBEC3G expression plasmid. Empty vector was used to maintain the same amount of total transfected DNA. The volumes used for infections were chosen to yield MOI 0.3 with APOBEC3G-minus virus. The Gag carried of each virus is indicated.

Figure 16


A
matrix
P12
capsid
nucleocapsid
protease
CERV2GAGconsensus
CERV2_MLVGAG2
MLVGAG
B
Viral Titer (IU/mL)
1.E+07
1.E+06
1.E+05
1.E+04
1.E+03
1.E+02
1.E+01
1.E+00
CERV2Gag-MLVPol
CERV2/MLVGag-MLVPOL
MLVGagPol
C
Infection (% of control)
100
10
1
CERV2
MLV
0
100
200
300
400
human APOBEC3G (ng)

Although CERV2 hypermutation only occurred in a dinucleotide context that was indicative of APOBEC3G activity (i.e. GG), RhERV2 hypermutation also occurred in a dinucleotide context that is targeted by several other APOBEC3 proteins (i.e. GA). Therefore, we asked if human APOBEC proteins other than APOBEC3G are capable of potently and specifically restricting CERV2, which would suggest that they could have protected human ancestors from infection by CERV2. Restriction of CERV2 Gag virions by most APOBEC proteins tested, though, was non-existent compared to APOBEC3G (Figure 18). It should be noted that editing of ApoB mRNA by APOBEC1 involves the APOBEC1 complementation factor (ACF) (273, 274) and perhaps inclusion of ACF could alter the results presented here. In contrast to the lack of activity observed in most APOBEC proteins tested, human APOBEC3F and APOBEC3A were able to moderately restrict CERV2 infection (two- and three-fold, respectively) but none of the restriction activities observed showed specificity for CERV2Gag virions over those containing MLV Gag. Although the data presented here indicate that APOBEC3 proteins were likely able to restrict CERV2 replication, there is no direct evidence that these restrictions were responsible for protecting the human lineage from CERV2 infection.

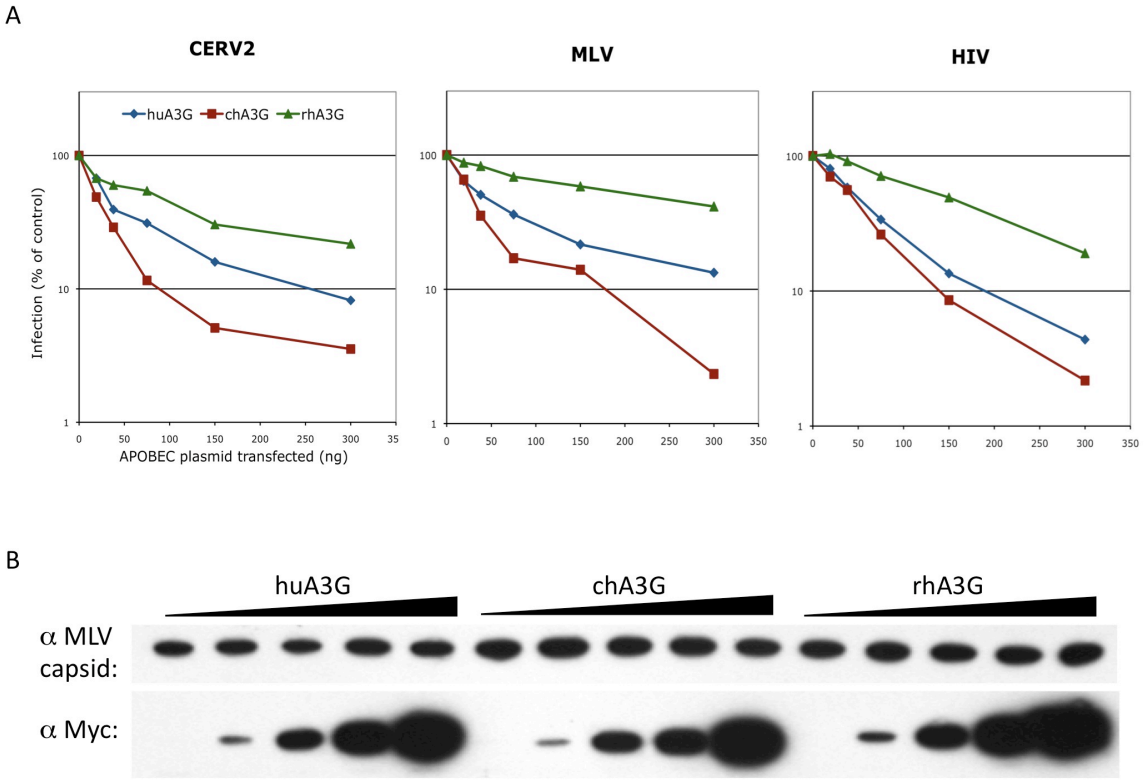


Figure 17. Restriction of CERV2 Gag-containing virions by human, chimpanzee, or rhesus macaque APOBEC3G. A) Titration of APOBEC3G was done as described in Figure 16C except myc-tagged APOBEC3G genes from three different species were used as indicated. Additionally, HIV virions were generated by cotransfection of VSVg, NL GagPol, and CSGW. B) Following harvesting of the supernatent used for infections in (A), 293T cells transfected with MLV GagPol were lysed in SDS-PAGE loading buffer for western blots probed with rabbit anti-MLV capsid or mouse anti-Myc antibodies to the detect the 46kDa APOBEC3G proteins.

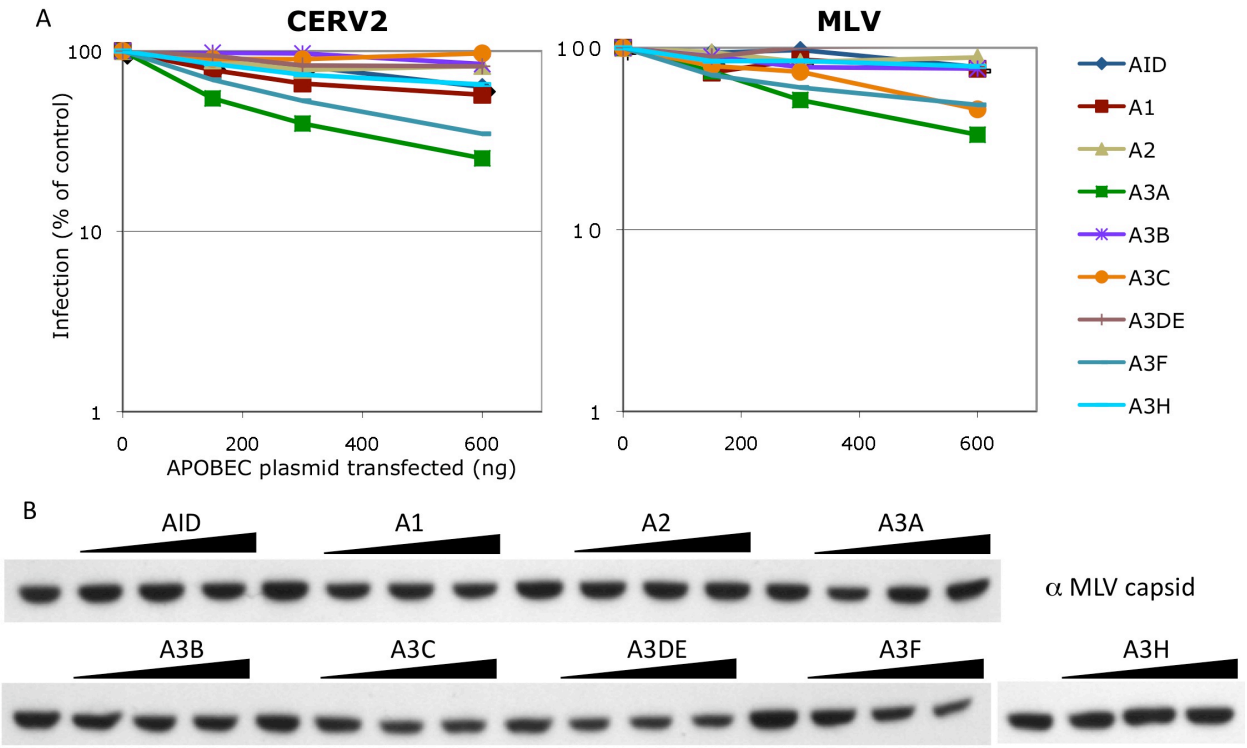


Figure 18. Lack of potent restriction of CERV2Gag-MLVPol by human APOBEC proteins other than APOBEC3G. A) The indicated human APOBECs were titrated as in Figure 17, except a maximum of 600ng (rather than 300ng) APOBEC expression plasmid was used. B) Lysates from MLV virus-producing cells used in (A) were used for western blot with a rabbit anti-MLV capsid antibody.

Chapter IV. The impact of APOBEC3 proteins on HIV-1 evolution in vitro

HIV-1 was used to investigate the potential impact of APOBEC3-induced mutation on a replicating viral population. The interaction between HIV-1 and APOBEC3 proteins is complex. Although its Vif protein protects HIV-1 from restriction by human APOBEC3 proteins, hypermutated proviruses are abundant in infected patients (192, 193, 197). It is currently unclear if these variants impose a significant burden to overall fitness of the viral population, thus constituting a deleterious genetic load. Another possibility, and one that is quite the opposite of the first, is that they can contribute an increase in variability that could benefit a rapidly evolving population, perhaps through recombination with fit viruses. Or, perhaps hypermutated proviruses neither positively nor negatively affect HIV-1 evolution due to debilitating mutational scars that have left their gene products entirely inactive and their genomes too unrecognizable to even be packaged by competent viruses. To explore these possibilities, we observed the accumulation of APOBEC3-induced diversity during HIV-1 replication *in vitro* and the resulting viral populations were tested for their ability to evolve under selective pressures.

Replication of HIV-1 delta Vif in permissive and semi-permissive cell lines

In order to assess the impact that APOBEC3-induced mutation has on a replicating viral population in terms of both its overall fitness and adaptive potential, HIV-1 or an HIV-1 lacking Vif expression (ΔVif) were allowed to replicate in human T-

cell lines. Both viruses expressed GFP in place of NEF (to track the spread of virus through the culture) and also contained mutations that prevented Vpr expression. The latter was used to preempt a loss of Vpr that can occur during long-term culture and can enhance virus replication (275). Two different cell lines, MT2 and CEMSS, both allowed replication of ΔVif virus but the Vif phenotype displayed in the two cell lines differed (Figure 19). The replicative fitness of ΔVif in CEMSS cells, as predicted by the literature, was identical to that of the virus carrying intact Vif (Vif+) (152, 160). The genetic screen that identified APOBEC3G as a host factor antagonized by Vif made use of this permissive phenotype of CEMSS (152). Replication of ΔVif in MT2 cells, on the other hand, was restricted relative to Vif+ but not but not enough to prevent a spreading infection that overtook the culture within five days- a two-day delay compared to Vif+. A Partial restriction of ΔVif replication has also been observed in activated PBMC and, therefore, the ΔVif phenotype observed in MT2 cells may be more realistic than that displayed in restrictive cell lines that entirely extinguish ΔVif replication (211).

The cytopathicity of the ΔVif and Vif+ viruses also differed. To maintain the MT2 culture infected with Vif+ (MT2/Vif+), fresh cells were added every three days beginning four days post-infection to counteract frequent cell death (the ratio of fresh to infected cells was approximately 1:4). In contrast, maintenance by the addition of healthy cells was not necessary in MT2/ΔVif cultures where cell death was rarely observed. The relative robustness of MT2 cells infected with ΔVif may be due to a lower level of infectious virus in the culture. A titration of cell culture supernatant collected 15 days post-infection (dpi) was carried out by infecting MT2 cells overnight before the

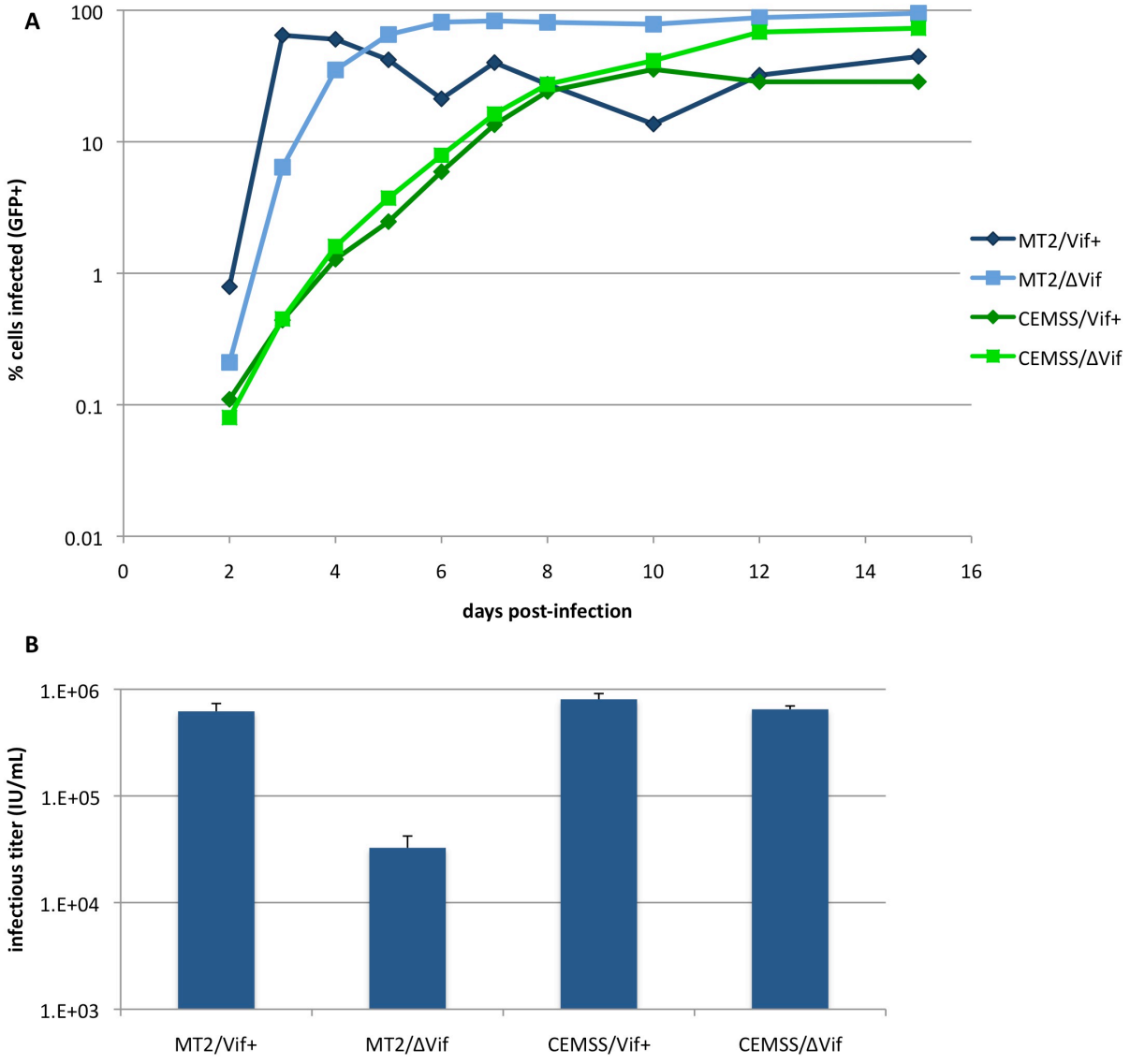


Figure 19. Replication of HIV with or without Vif in MT2 or CEMSS cells. A) 10^6 MT2 or CEMSS cells were infected at MOI 0.001 with HIV-1 NL4.3/HxB/GFP delta Vpr with Vif (Vif+) or deficient in Vif (ΔVif). Aliquots of cells were fixed periodically for detection of GFP-positive cells by FACs analysis. <0.1% GFP-positive (or <5 cells) is considered background based on measurements of un-infected cells. B) Supernatants were collected from cultures described in (A) at 15 dpi and titrated on 10^4 MT2 cells. Infections proceeded overnight before addition of dextran sulfate, followed by fixation at two days post-infection. Infectious titer was derived from the percentage of cells positive for GFP.

addition of dextran sulfate to prevent multiple replication cycles. The titer of infectious virus in supernatent from MT2/ΔVif was 20-fold lower than that of MT2/Vif+ at 15 dpi (Figure 19B). It should be noted that these infections included restriction by any APOBEC3 proteins that may have been packaged into virions from the previous culture.

The behavior of infected CEMSS was different from that observed in infected MT2. The infected CEMSS displayed frequent syncytia rather than cell death and these effects were similar between CEMSS cultures infected with either Vif+ or ΔVif. To prevent fusions from occurring in the vast majority of cells, fresh CEMSS were periodically added in equal amounts to both of the infected cultures. Still, after ten days of replication, the CEMSS culture infected with Vif+ had a lower percentage of cells that were GFP-positive compared to CEMSS infected with ΔVif, which may have been a reflection of weak GFP expression in unhealthy cells (Figure 19A). There was no apparent difference, though, in the infectious titer of the supernatent in each culture at 15 dpi (Figure 19B). These results suggest that Vif may be toxic to the cells and could contribute to the cytopathic effect of HIV replication in some cell types.

Replication of HIV-1 with or without Vif in restrictive cells presumably resulted in viral populations that differed greatly in their sequence diversity due to APOBEC3-induced mutation. It remained possible, though, that ΔVif was able to replicate efficiently in MT2 cells because of a weak penetration of APOBEC3 activity that allowed the predominance of individual viruses entirely untouched by deamination over a minority population of severely hypermutated viruses. To determine if APOBEC3-induced sequence diversity could be observed following a period of continual viral replication, ΔVif or Vif+ were allowed to replicate in MT2 or CEMSS cells for 15 days

before cells and supernatent were collected. Genomic DNA was purified from cells and RNA was extracted from pelleted virions for cDNA synthesis. A 720-nucleotide locus covering the polymerase domain of reverse transcriptase was amplified from both proviral DNA in cellular genomes as well as cDNA made from virion-packaged viral RNA. The resulting PCR products were cloned and twelve clones from each were sequenced and compared to the wildtype sequence (Figure 20). All twelve proviral sequences from MT2/ΔVif contained varying degrees of mutations that were likely caused by APOBEC3-catalyzed deamination (Figure 20A). These G to A substitutions in GG or GA dinucleotide contexts varied in quantity from six to 38 per amplicon. Surprisingly, APOBEC3-associated mutations were also detected in ΔVif proviral DNA from non-restrictive CEMSS cells, albeit at levels that were much lower than those observed in ΔVif proviral sequences amplified from MT2 cells (25 mutations in one CEMSS/ΔVif sequence and 1-12 mutations in six other clones out of 12 examined). This suggests that levels of APOBEC3 proteins are present in CEMSS cells that are able to accomplish a degree of catalytic deamination but are not sufficient to restrict virus replication. In contrast, only a single G to A mutation within a GA or GG context was found in proviral sequences from the MT2/Vif+ culture and none were found in sequences from CEMSS/Vif+. In general, other types of mutations were found in a minority of the sequences analyzed at a maximum of only one substitution per sequence.

An analysis of virion-packaged sequences revealed a different degree of GG to AG and GA to AA mutations compared to those observed in proviruses (Figure 20B). Although APOBEC3-induced mutations were observed again in the MT2/ΔVif culture, they were not found in every sequence (8/12) and, where present, there were only 2-10

mutations per sequence. A decreased degree of hypermutation in packaged viral RNA relative to integrated provirus may represent a purification of sequences that are either not mutated or only moderately mutated by APOBEC3. Such selection could be explained if severe hypermutation resulted in impaired LTR promoter function, the degradation of RNAs that contain frequent stop codons by the nonsense-mediated decay pathway, or the crippling of the RNA structures responsible for packaging into virions by nucleocapsid. Interestingly, there were three virion sequences from MT2 cultures infected with Vif+ virus and two from CEMSS infected with ΔVif virus that contained a small number of APOBEC3-associated mutations (1-3 mutations/sequence). In spite of the small sample size of sequences in these analyses, it is clear that the population of ΔVif viruses in MT2 cells contained APOBEC3-driven diversity that was represented at the level of viral genomic RNA. This suggests that an encounter of a viral genome with APOBEC3 is not necessarily a dead end but, rather, hypermutated sequence may be carried by virions during subsequent rounds of replication. Furthermore, it *cannot* be concluded that either Vif+ virus replicating in ΔVif-restricting cells or ΔVif virus in non-restrictive cells are entirely devoid of APOBEC3-induced mutation. The experiments described below explored whether this varying degree of diversity can provide a benefit or a burden to evolving viral populations.

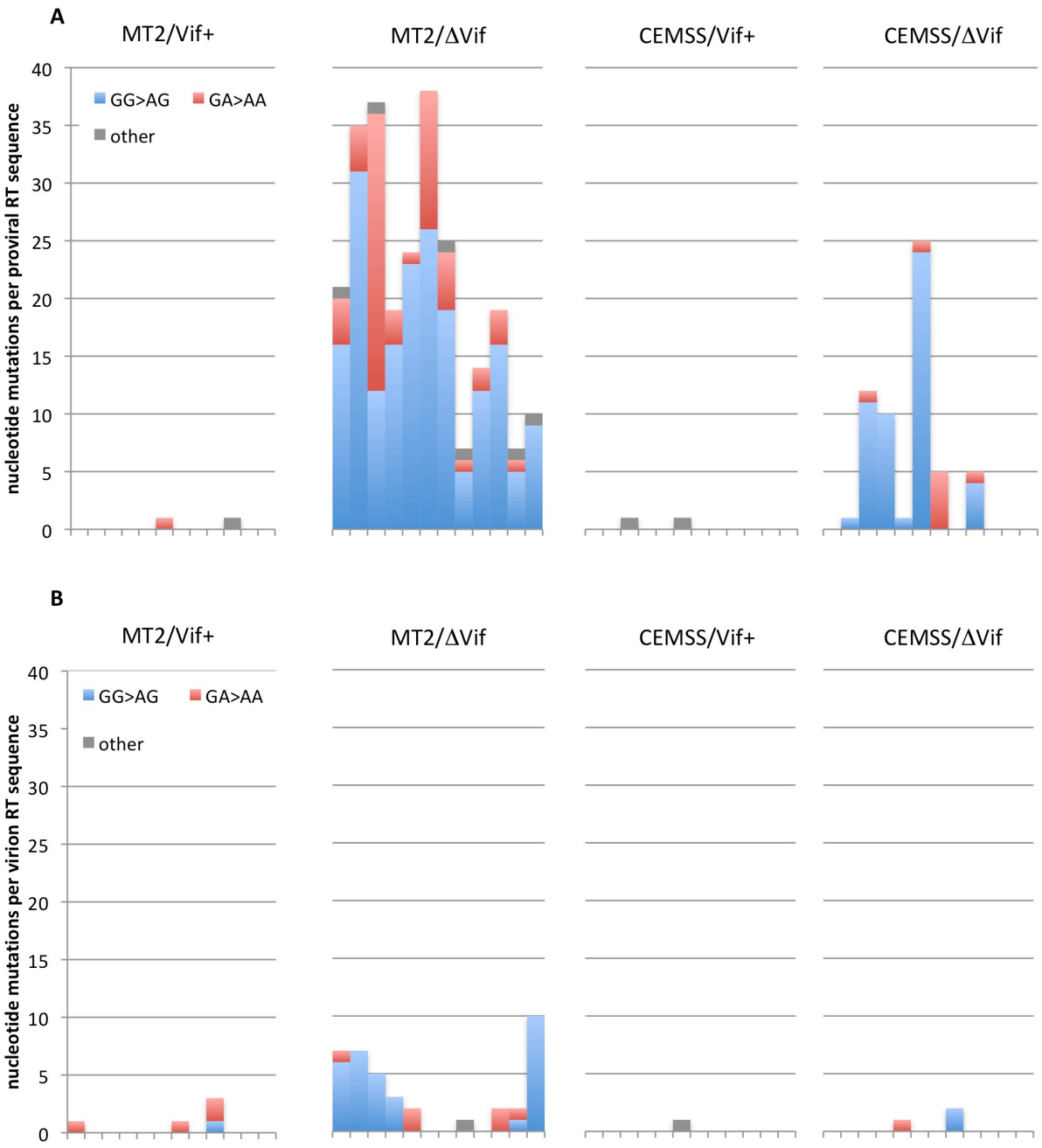


Figure 20. Detection of APOBEC3-induced mutations. A and B) DNA encoding the polymerase domain of HIV reverse transcriptase was amplified, cloned, and sequenced. The number of nucleotide substitutions relative to the wildtype sequence is shown for each sequence. Putative APOBEC3-induced mutations are indicated as G to A changes that immediately proceeded a G or A residue (blue or red, respectively) and all other nucleotide substitutions were pooled (grey). Sequences were divided according to the inoculating virus (Vif+ or ΔVif) and the cells in which it replicated (MT2 or CEMSS). Templates for PCR were either cellular DNA preparations (A) or cDNA derived from pelleted virions (B).

Selection of drug resistance in Vif+ virus

We wished to determine if the potential for HIV-1 to acquire drug resistance differed according to the degree of APOBEC3-driven diversity. First, ΔVif or Vif+ viruses that had been grown in the absence of selection in MT2 cells (MT2/ΔVif and MT2/Vif+ viruses) were used to infect fresh MT2 cells at MOI 0.001. AZT, nevirapine, or a combination of the two was added to the infected cultures on the following day (Figure 22). The drugs were chosen because they are known to induce the emergence of mutations that could potentially be caused by APOBEC3 activity (Table 1). AZT and nevirapine were further selected because they offered a range of concentrations in which virus spread in MT2 cells was inhibited but not entirely extinguished (Figure 21). In contrast, inhibition with nelfinavir jumped from slight delays in viral spread to a complete inhibition with only a two-fold change in drug concentration; while high concentrations of 3TC, at best, caused a two-day delay in the complete spread of virus through the culture.

Table 1. Drug resistance mutations that could result from APOBEC3 activity

Nevirapine	AZT	3TC	Nelfinavir
G190E	D67N	M41I	D30N
G196E	R211K	M184I	M46I
M230I		R211K	R57K
			G73S

The amino acid substitutions listed are associated with resistance to the indicated drug and are induced by GG to GA or GA to AA nucleotide mutations. The NL4.3 GagPol sequence was used to determine the dinucleotide context of the mutations. Based on data compiled in the Los Alamos National Laboratory and Stanford University HIV Drug Resistance Databases and other cited references (211, 276-294).

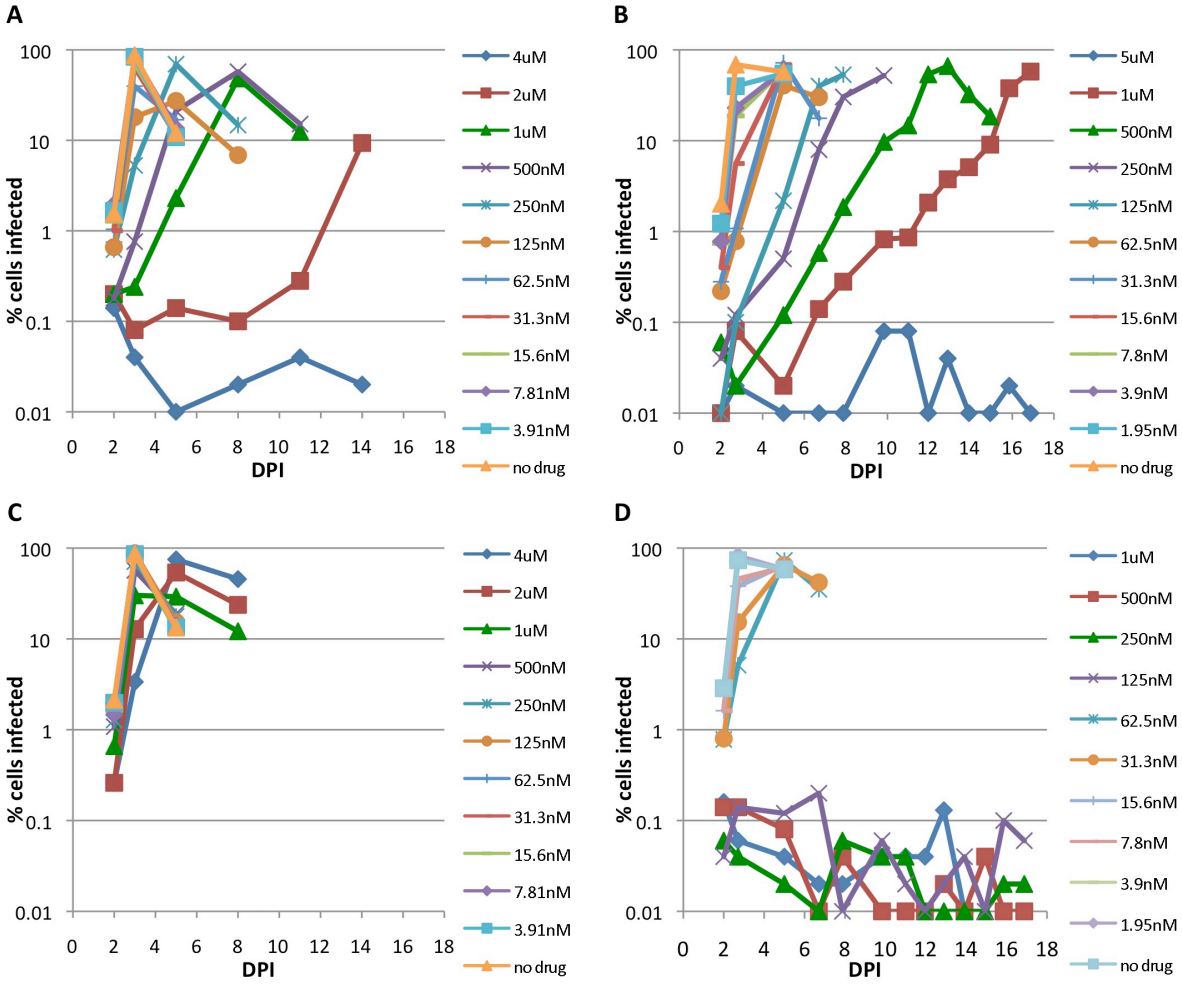


Figure 21. Inhibition of spreading infection in MT2 cells by antiretroviral drugs. 10^6 MT2 cells were infected overnight with Vif+ and then split into media containing nevirapine (A), AZT (B), 3TC (C), or nelfinavir (D) at the indicated concentrations. Cells were allowed to grow on 24-well plates with a 1:5 passage into media containing fresh drug every four days. 100µL of cells were fixed at several time-points for quantitation of GFP-expressing (or infected) cells.

The Vif+ virus overtook the MT2 culture within 10 days in the presence of AZT and it was also able to rapidly spread through a nevirapine-containing culture after an initial decline in the number of infected cells (Figure 22A, B). Additionally, it took MT2/Vif+ about a month to emerge under suppression with a combination of both drugs (Figure 22C). In contrast, MT2/ΔVif was not able to overcome any of these three challenges. A titration of nevirapine or AZT in TZM cells during infection with the putative drug-resistant viruses revealed that nevirapine resistance had, indeed, been acquired by MT2/Vif+ during both single and dual-drug selections; but neither AZT alone nor AZT in combination with nevirapine had induced any drug resistance (Figure 23). Inhibition with AZT at the chosen concentration was apparently only sufficient to delay the rapid spread of Vif+, while nevirapine was potent enough to elicit the emergence of drug resistance mutation(s). Amplification, cloning, and sequencing of reverse transcriptase polymerase domain sequences from proviruses at the end of selection with nevirapine revealed a variety of mutations associated with nevirapine resistance in 8/10 sequences (Figure 24). One of these amino acid substitutions, M230I, was the result of a GG to AG mutation. M230I was found in two sequences: one only contained a single point mutation and the other contained 24 other APOBEC3-induced amino acid mutations, including three stop codons. The hypermutant also contained another predicted nevirapine resistance mutation, Y188C, which was not caused by a G to A substitution and was shared by three other sequences. It is possible that during the course of its escape, this individual accumulated one or both beneficial mutations before a detrimental hypermutation by APOBEC3.

The lack of drug resistance evolution in ΔVif virus could be due to a continual restriction by APOBEC3 proteins in MT2 cells during selection rather than a deleterious effect of pre-existing genetic loads. Perhaps restriction, although not enough to prevent the spread of ΔVif through a culture that lacked drugs, severely decreased the odds of an emergence by rare escape mutants. To address this, selection with nevirapine was repeated using CEMSS cells where a lack of Vif did not result in lower replication kinetics. In this experiment, viruses that were grown for 15 days without selection in MT2 cells (termed MT2/Vif+ and MT2/ΔVif viruses) as well as viruses grown in CEMSS cells (termed CEMSS/Vif+ and CEMSS/ΔVif viruses) were used to infect fresh CEMSS at an equivalent MOI of 0.001. Nevirapine was added three days after infection at a concentration that was capable of inhibiting a spreading infection in CEMSS cells (Figure 25A). It should be noted that even in the absence of drug, MT2/ΔVif displayed a slower replication kinetic than the other three viruses- perhaps the effect of a deleterious load of mutations (Figure 25B). Three weeks after infection, only CEMSS/Vif+ was able to overcome inhibition by nevirapine; which, out of the four populations used to initiate the selection, was the most likely to be devoid of any APOBEC3-induced mutations. Therefore, these experiments demonstrate that immense diversity generated by APOBEC3 in the absence of Vif does not necessarily make a productive contribution to drug resistance evolution. It is possible that such potential can be outweighed by an associated genetic load. Nevertheless, the occurrence of M230I in MT2/Vif+ following selection with nevirapine, suggests that it is possible for low levels of APOBEC3-induced mutation in Vif-positive viruses to productively contribute to evolution. It is worth noting that in order to attain a low MOI before selection began, a bottleneck was imposed

on the viral population and it is possible that this resulted in the loss of some potentially beneficial APOBEC3-induced mutations. Perhaps the use of larger initial population sizes prior to selection could allow the inclusion of broader diversity and, thus, reveal greater evolutionary potentials.

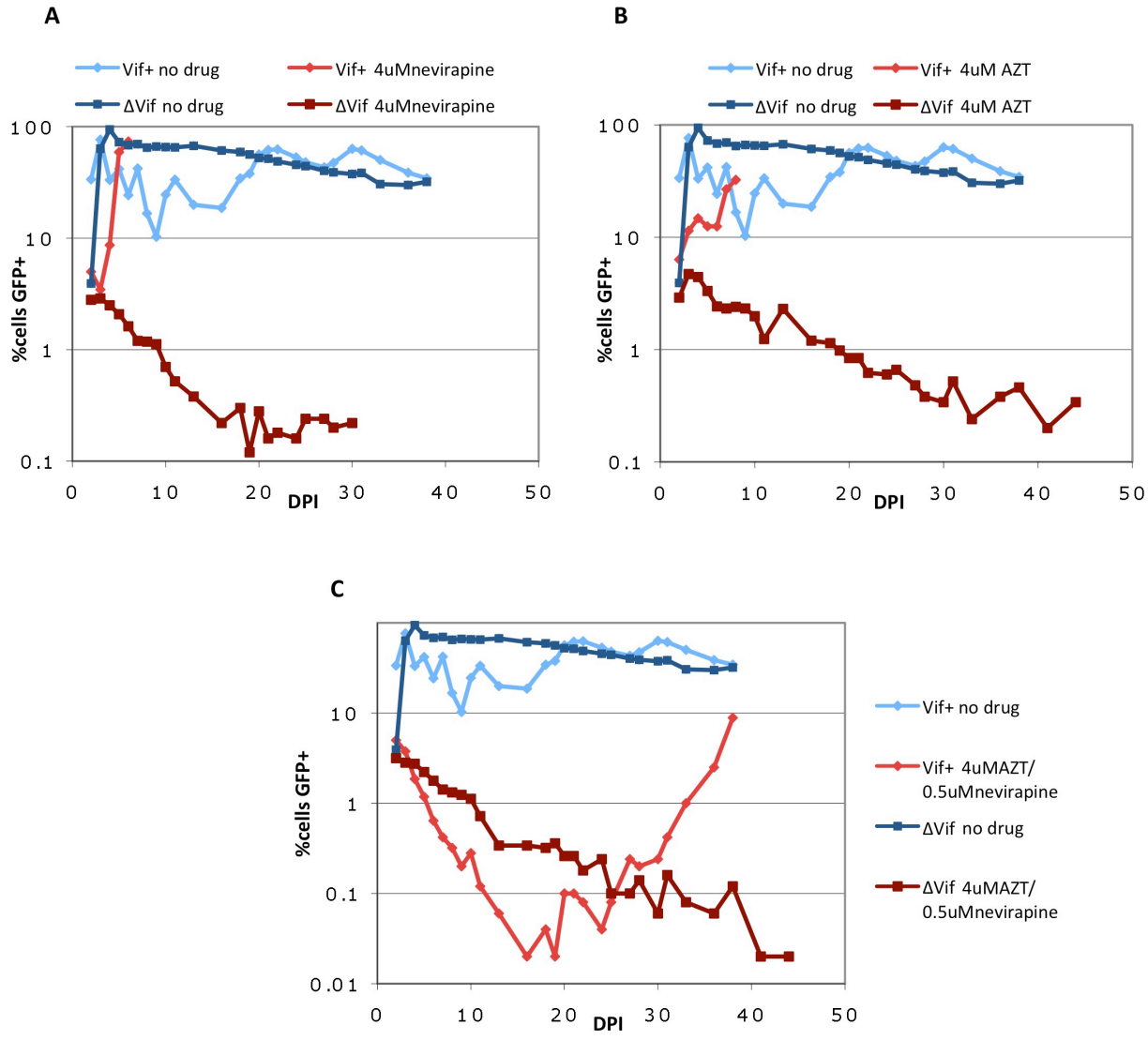


Figure 22. Single and dual-drug selection in MT2 cells. Vif+ or ΔVif viruses were grown in MT2 cells for 15 days (Figure 19), and then used to infect fresh MT2 at MOI 0.01. One day after infection, cells were split into plain media, 4μM nevirapine (A), 4μM AZT (B), or a combination of 4μMAZT and 0.5μM nevirapine. Growth curves are plotted as percent infected cells (expressing GFP) verses days post-infection.

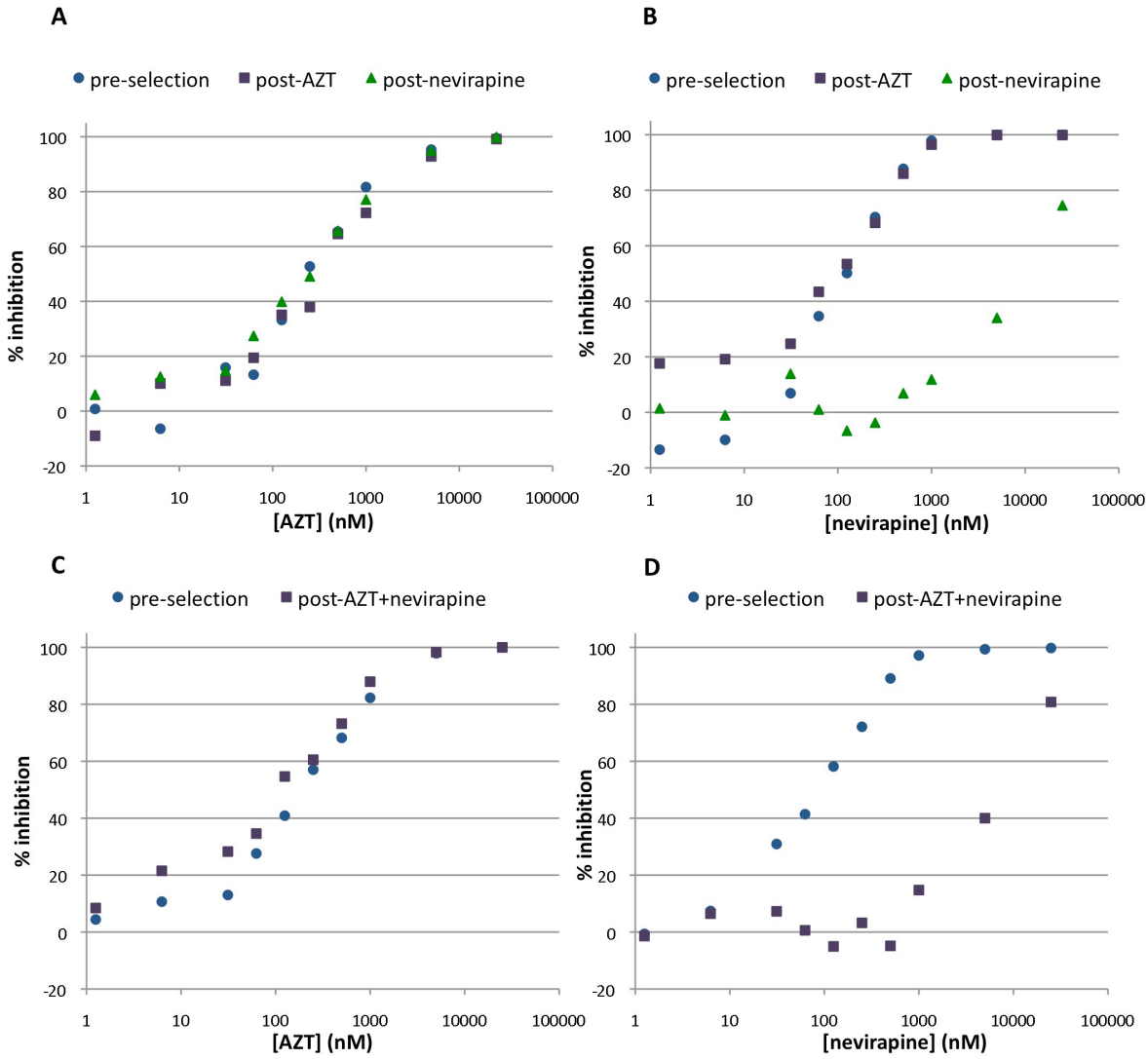
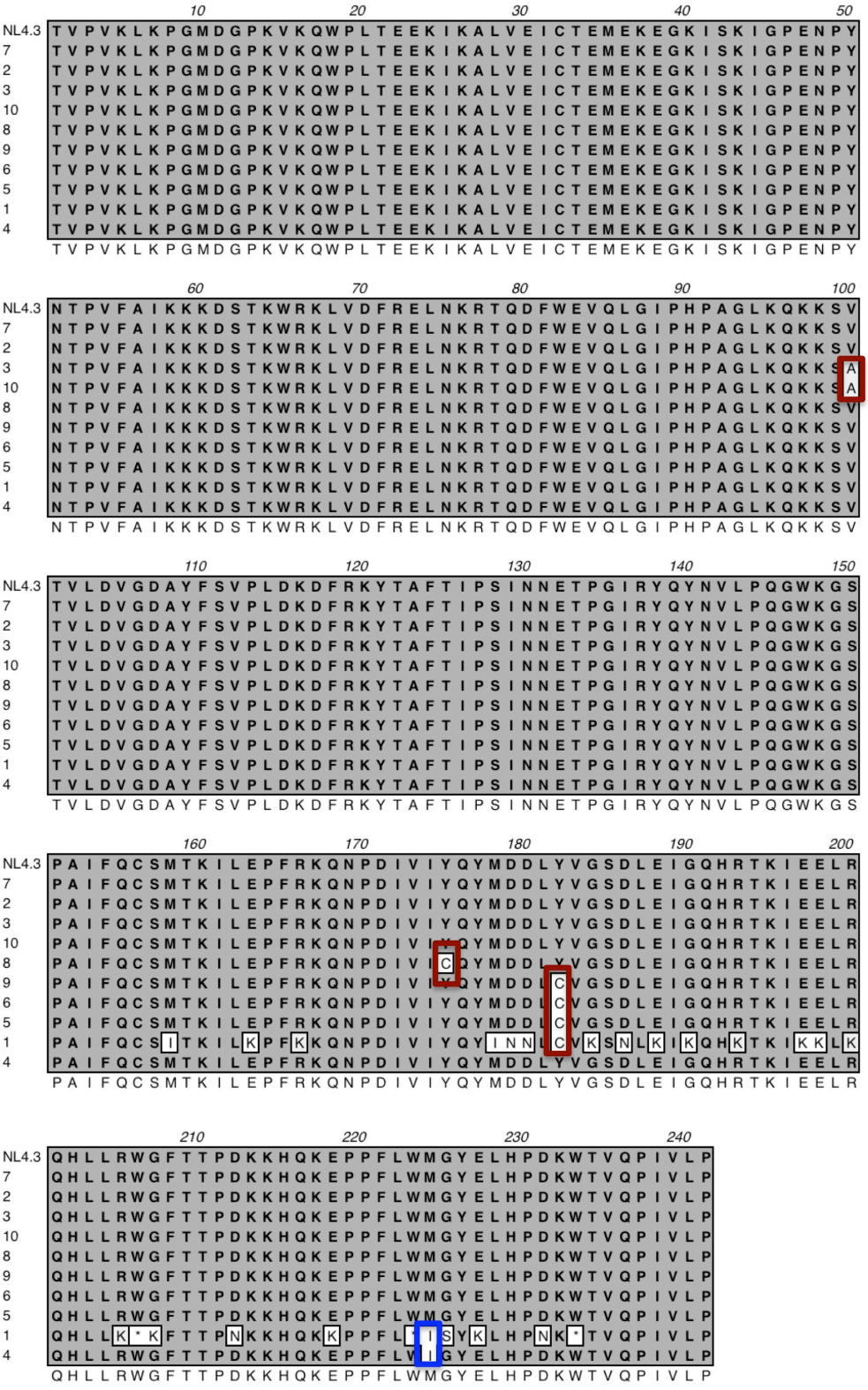


Figure 23. Evolution of nevirapine but not AZT resistance. TZM cells were infected at MOI 0.3 in media containing increasing concentrations of AZT (A and C) or nevirapine (B and D). Cells were infected with either the Vif+ virus used to initiate selection experiments or resultant Vif+ viruses after spread through MT2 cells in the presence of the indicated drug (A and B) or drugs (C and D). TZM infection was quantified two d.p.i. in relative luminescence units by a luciferase activity assay. Inhibition relative to a minus-drug control (%) is plotted verses drug concentration.

Figure 24. Nevirapine resistance mutations. Following selection with 4μM nevirapine, proviral sequences encoding the reverse transcriptase polymerase domain were amplified from MT2 cellular DNA, cloned and sequenced. An alignment of the amino acid sequence from 10 clones and the wildtype NL4.3 sequence is shown. Primer-binding sites, including the first six codons of reverse transcriptase, were excluded. Sequences with common mutations were ordered together. Amino acid changes predicted to confer a decrease in sensitivity to nevirapine are marked with boxes (286, 287, 290, 291, 293, 294) and one of these (marked in blue) was caused by a G to A mutation in a GG dinucleotide context.

Figure 24


10 20 30 40 50
NL4.3 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
7 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
2 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
3 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
10 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
8 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
9 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
6 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
5 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
1 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
4 TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
TVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPY
60 70 80 90 100
NL4.3 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
7 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
2 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
3 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSA
10 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSA
8 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
9 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
6 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
5 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
1 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
4 NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
NTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKQKKSV
110 120 130 140 150
NL4.3 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
7 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
2 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
3 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
10 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
8 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
9 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
6 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
5 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
1 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
4 TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
TVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGS
160 170 180 190 200
NL4.3 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
7 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
2 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
3 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
10 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
8 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
9 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
6 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
5 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
1 PAIFQCSITKILKPFKKQNPDIVIYQYINNLYVKSNLKIKQHKTKIKKLK
4 PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
PAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELR
210 220 230 240
NL4.3 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
7 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
2 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
3 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
10 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
8 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
9 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
6 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
5 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
1 QHLLKSKFTTPNKKHQKKPPFLTISYKLHPNK*TVQPIVLP
4 QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP
QHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP

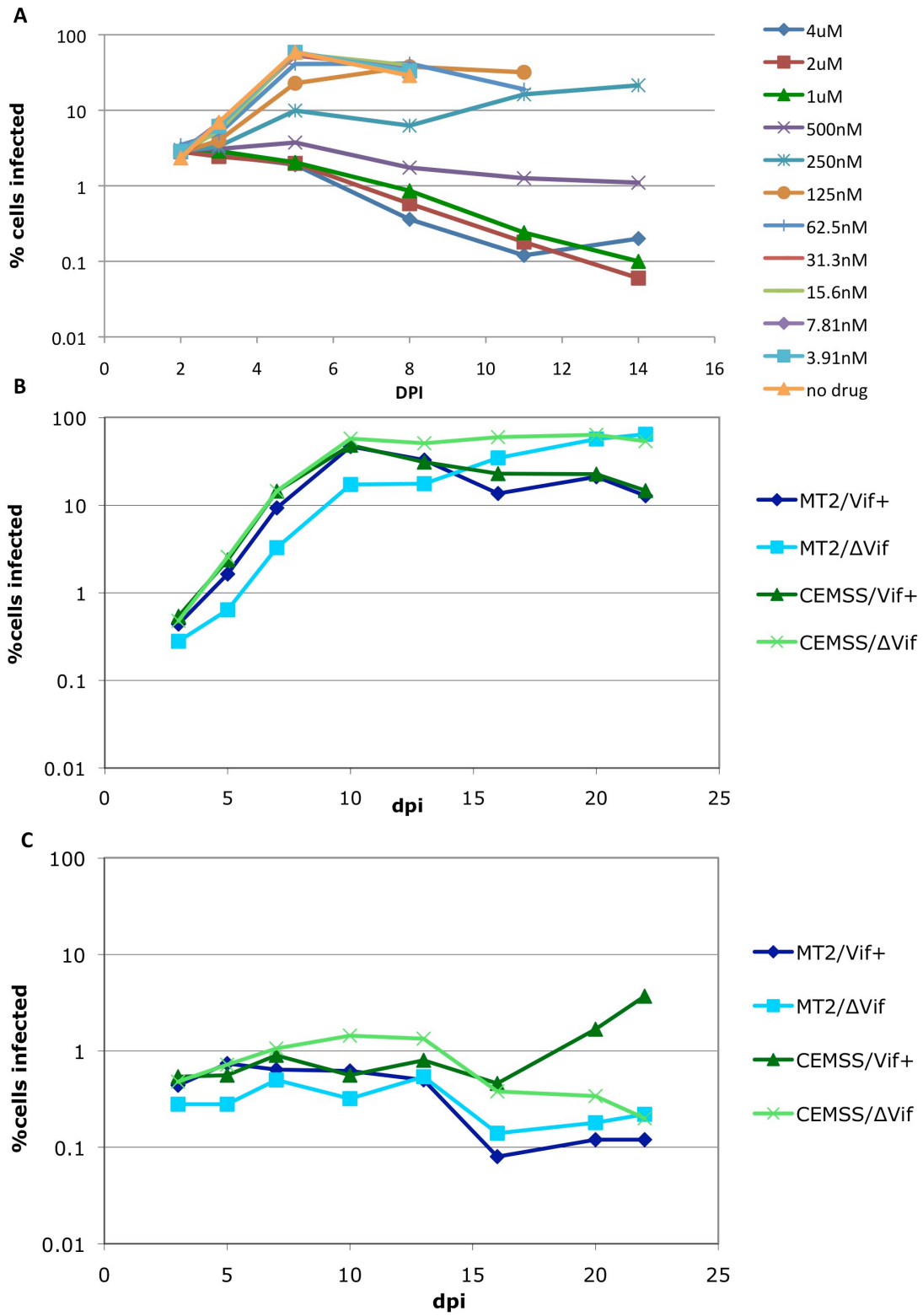


Figure 25. Selection with Nevirapine in CEMSS cells. A) Titration of nevirapine in infected CEMSS was done as described in Figure 21 for drug titrations in MT2 cells. B and C) Vif+ or ΔVif grown 15 days in MT2 or CEMSS cells were used to infect CEMSS cells at MOI 0.001. Three days later, cells were split into plain media (B) or media containing 500nM nevirapine (C).

Selection of an escape mutant using patient plasma with neutralizing activity

In addition to *in vitro* drug-resistance evolution experiments, we also attempted to evolve escape mutants under selection by antibodies present in broadly neutralizing sera. Aside from providing a further examination of the evolutionary potential of APOBEC3-mutated viruses, this type of selection experiment could also potentially identify novel antigens targeted by neutralizing antibodies or more general determinants of neutralization sensitivity. Selective pressure was imposed by blood plasma containing broadly neutralizing activity that was taken from an HIV-1 infected, untreated individual with long-term low levels of viremia (i.e. a long-term non-progressor) (295). For this experiment, Vif+ virus was grown in MT2 cells and, therefore, only trace levels of APOBEC3-induced mutations were expected. One day after infection at MOI 0.001, MT2 cells were split into varying dilutions of neutralizing plasma. Every four days, cultures were split 1:5 into media containing fresh plasma. A 1:20 dilution of the plasma was initially able to prevent detectable virus replication for one week. Then, a steady spreading infection was apparent over an additional one-week period before a plateau persisted with 3-4% of the cells infected (Figure 26A). The resulting virus was titrated and used in a second round of selection alongside the original virus stock (Figure 26B). In spite of a slight fitness disadvantage, the selected virus was far less susceptible to neutralization than its un-selected counterpart and after approximately two weeks of additional selection, cells and supernatent were harvested for further analysis.

A titration of the patient plasma in a single-cycle infection demonstrated that two rounds of selection had produced a virus with decreased sensitivity to the plasma's neutralization activity (Figure 27A). Envelope sequences were amplified from cellular

DNA and ten clones were sequenced in order to identify putative escape mutations. A point mutation in variable loop three (V3) of GP120, K306E, was common to all sequenced clones (Figure 28). An envelope gene that did not contain any changes other than K306E was subcloned into the parental virus and a titration of neutralizing plasma demonstrated that this single mutation entirely recapitulated the decreased sensitivity observed in the bulk population of passaged virus (Figure 27A).

The V3 loop has previously been identified as a target of neutralizing antibody binding, but mutations to the V3 loop have also been implicated in conformational changes that shift the general neutralization sensitivity of GP120. Specifically, a gain of positive charges, particularly in residues 306 and 320, are thought to "open" the conformation of GP120, which allows a shift in co-receptor usage from CCR5 to CXCR4; but this conformational change is thought to also render the virus more neutralization sensitive (54, 55). Here, the starting virus was X4-tropic and it is possible that the K306E is not a specific antigen binding-site mutation that emerged due to escape from recognition by a particular antibody, but rather it may have caused a conformational change that was generally protective against a variety of antibodies. Although MT2 cells do not express CCR5, K306E was still tested for its ability to utilize it as a co-receptor to determine if the virus collaterally gained dual-tropism during neutralization escape. Infection of CHO cells expressing a humanized cyclin-T (to allow HIV transcription), human CD4, and either human CCR5 or human CXCR4 (259) showed that K306E was still only X4-tropic (Figure 27B). Nevertheless, the emergence of a residue associated with R5-tropism during neutralizing antibody escape, supports a model in which residues associated with X4-tropism are inherently more neutralization sensitive.

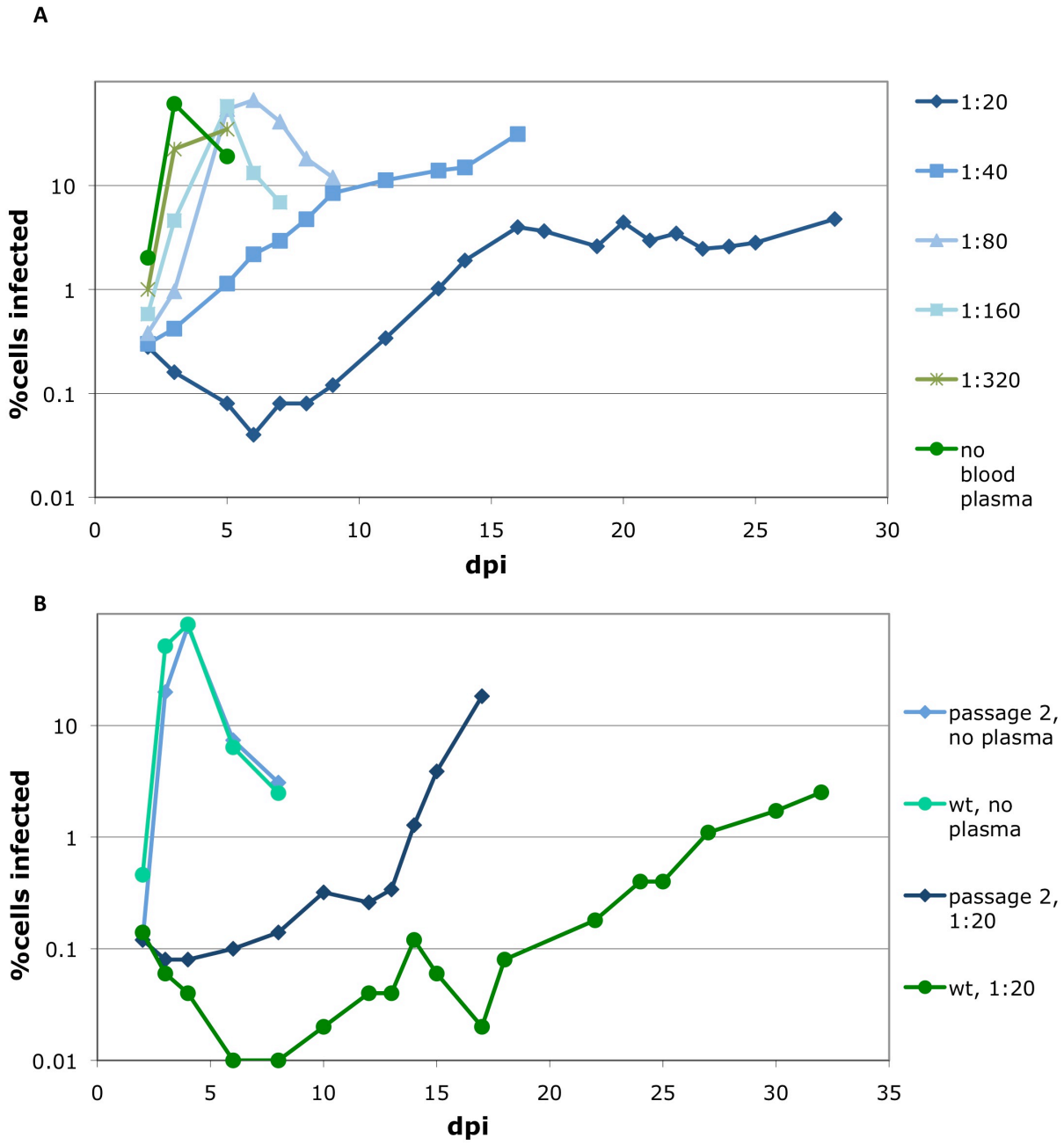


Figure 26. Selection using patient blood plasma with virus-neutralizing activity. A) MT2 cells were infected with clonal Vif+ virus at MOI 0.001. The following day, infected cells were split into plain media, or two-fold serial dilutions of patient plasma. Fresh plasma was added at the same dilution when cells were split every four days. B) Second Passage. MT2 cells were infected with virus selected in a 1:20 dilution of plasma (from A) or the same virus used to initiate the first passage, both at MOI 0.001. The next day, both infection were split into plain media or media containing a 1:20 patient blood plasma dilution.

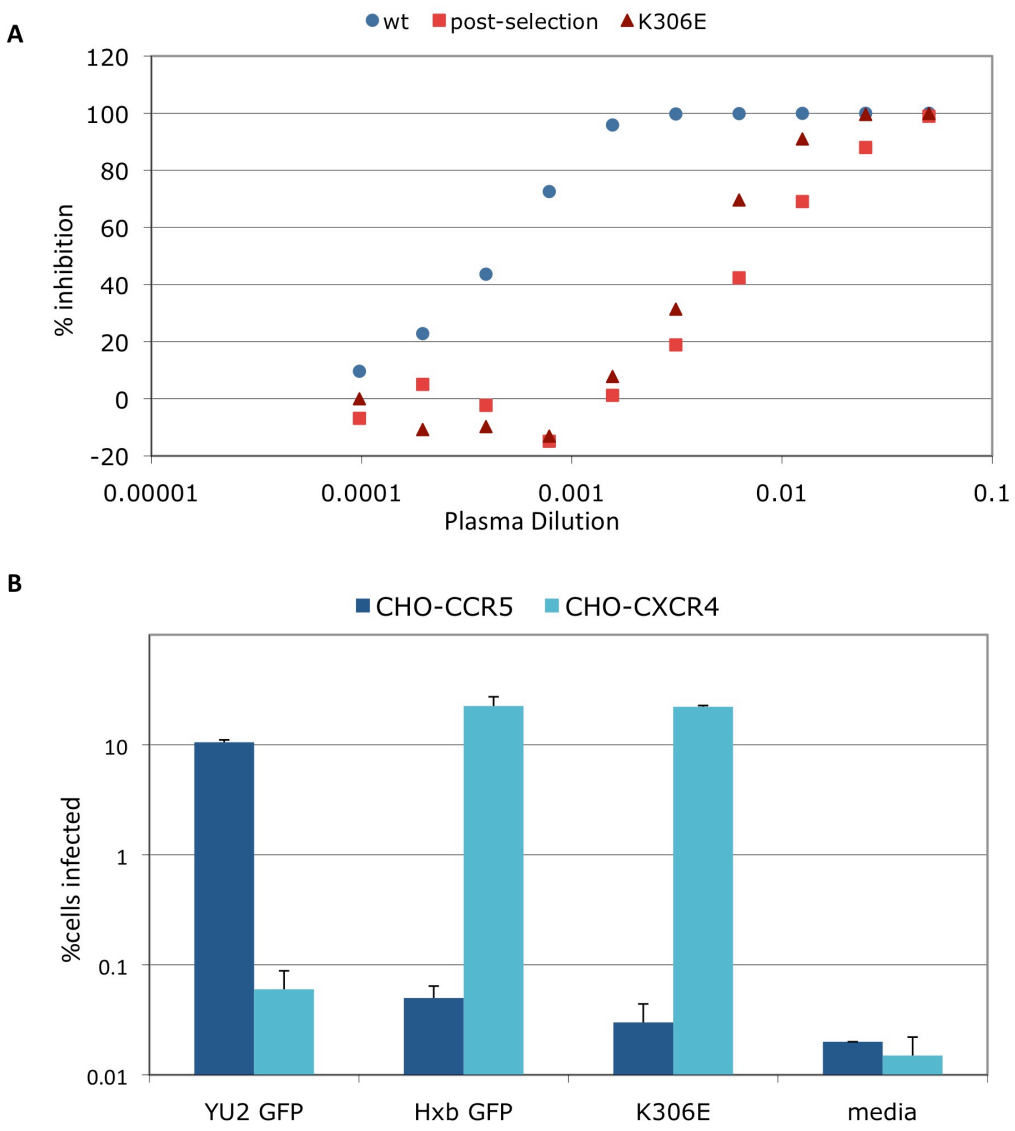


Figure 27. Escape mutants show decreased sensitivity to neutralization and remain X4-tropic. A) HIV-1 used to initiate plasma selection (wt), virus resulting from two rounds of selection with a 1:20 dilution of plasma (post-selection), or virus containing a lysine to glutamate mutation at envelope residue 306 (K306E) were used to infect TZM cells in the presence of neutralizing plasma at the indicated dilutions. As was done for drug titrations in TZM cells, infection was quantified two d.p.i. in relative luminescence units by a luciferase activity assay. Inhibition relative to a minus-plasma control (%) is plotted verses plasma dilution. B) HIV-1 GFP containing the R5-tropic YU2 envelope, the X4-tropic HXB-2 envelope, or HXB-2 envelope with the K306E mutation was used to infect 10^4 HIV-1 permissive CHO cell lines expressing either human CCR5 or human CXCR4. An average of two infections is shown for each.

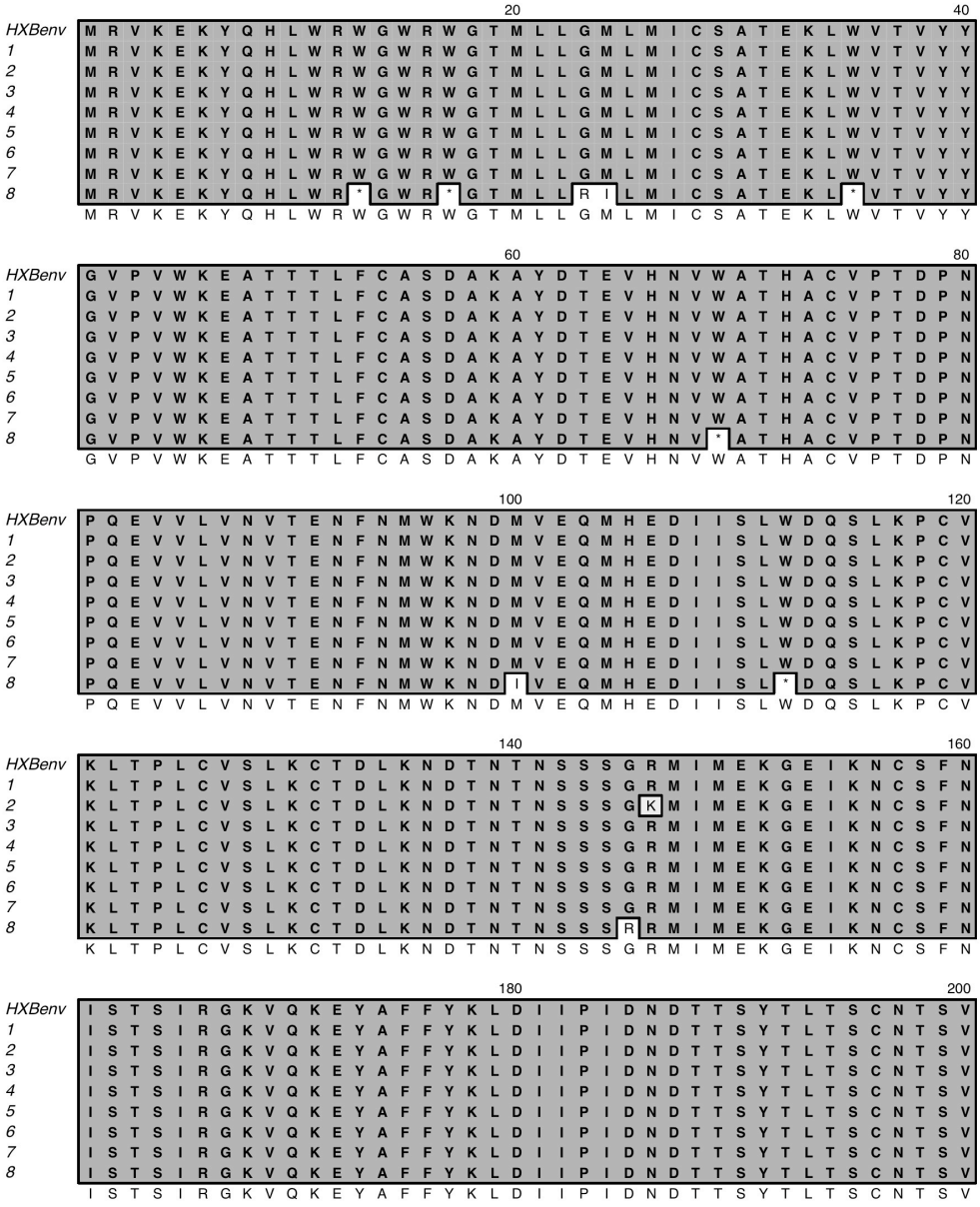


Figure 28. A V3 loop mutation was fixed into viral population during selection by neutralizing plasma. Proviral envelope sequences amplified from MT2 cells after two rounds of selection in neutralizing plasma were aligned with a confirmed wildtype HIV-1 HXB envelope sequence. Blue lines mark the boundaries of the V3 loop and the K306E mutation is indicated with an asterisk.

Figure 28 continued

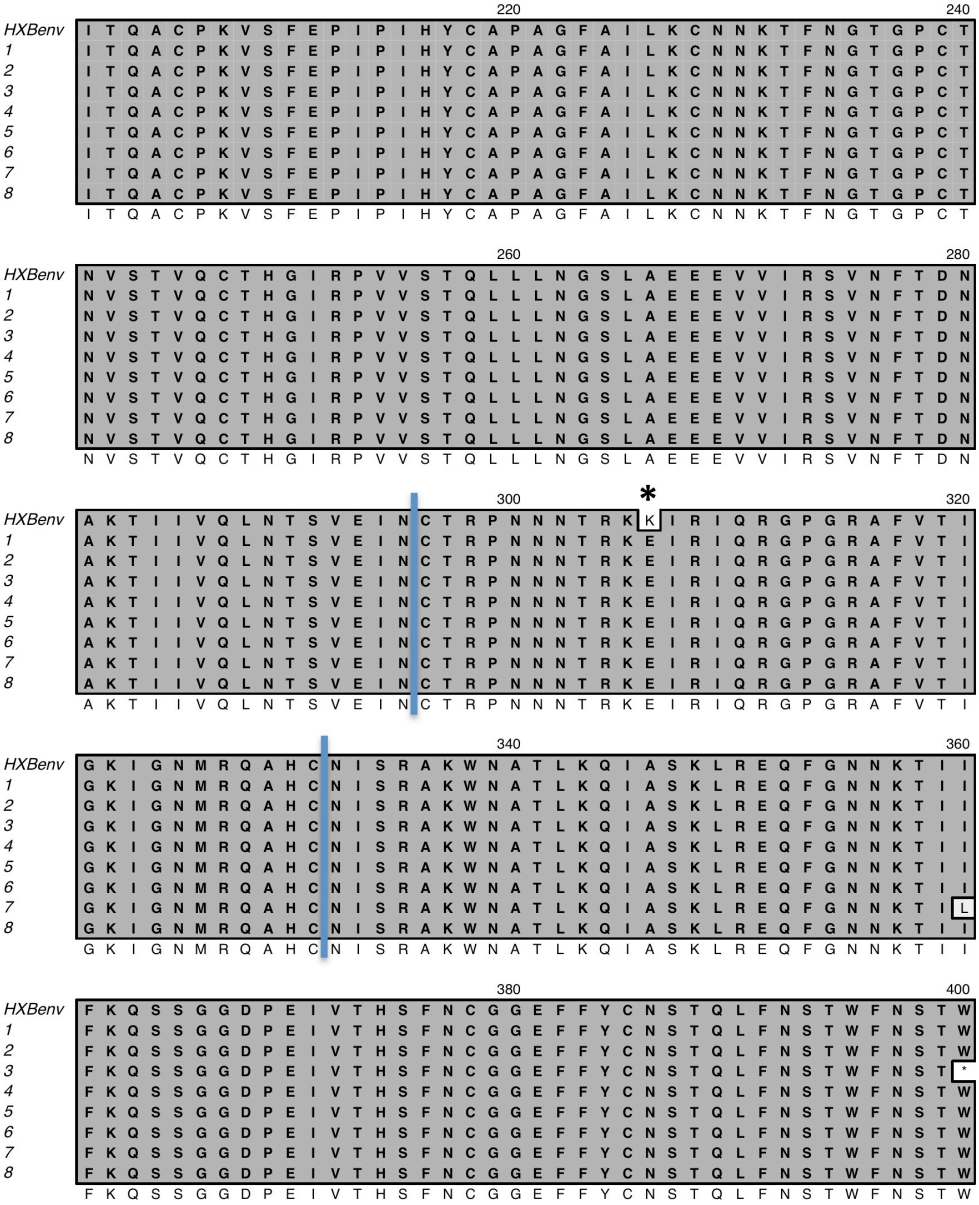


Figure 28 continued

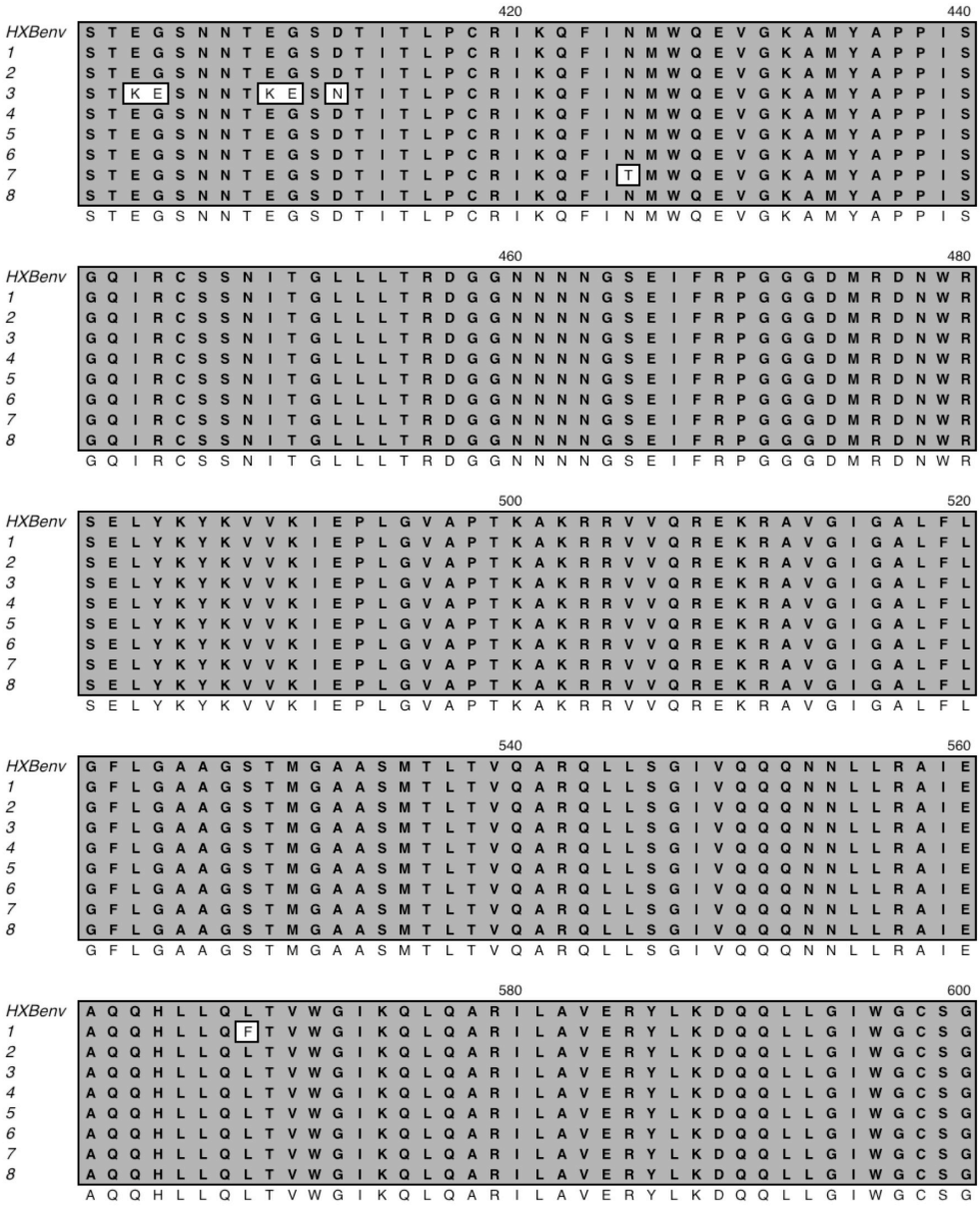
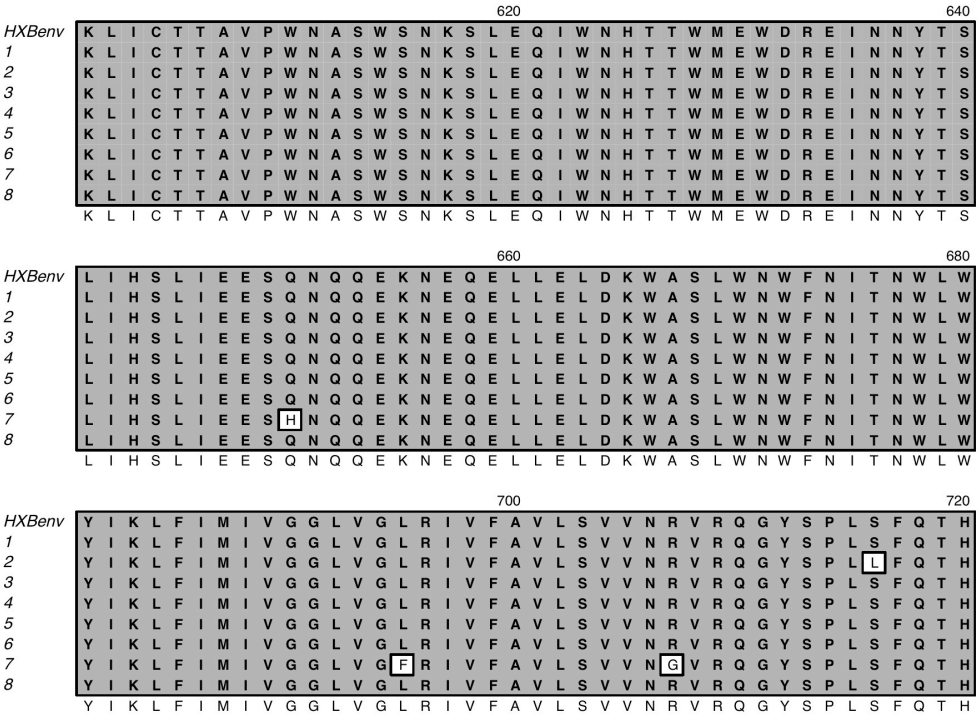


Figure 28 continued



Chapter V. Identification of the CERV2 cell surface receptor

While it remains possible that, as a means of defense, APOBEC3 proteins were able to assault CERV1 and CERV2, the exploitation of other host factors by these viruses would have been crucial to their successful infection of primate ancestors. In order to bind and enter cells, CERV1 and CERV2 had to hijack host proteins that normally reside on the cell surface to perform a task for the host. It is likely that the identity of these receptors determined the types of cells that CERV1 and CERV2 were able to enter and, perhaps, even the range of species they were capable of infecting. Therefore, to understand the ancient inter-species transmissions and endogenizations of CERV1 and CERV2, it is necessary to identify their cell-surface receptors.

In order to study the host range and receptor usage of CERV1 and CERV2 during the time that they replicated, we constructed consensus envelope genes from both virus families. Figure 29 shows an alignment of CERV2, RhERV2, and CERV1 consensus envelopes with moloney MLV envelope. As is common for gammaretroviral envelopes, a high degree of conservation was observed in the transmembrane protein (TM also known as p12E) compared to a divergent surface protein (SU also know as gp70). As expected, sequences were especially divergent in the regions that are responsible for receptor binding in MLV (variable regions A and B); with the exception of CERV2 and RhERV2, whose conservation in these regions may predict a common receptor usage.

Figure 29. Clustal alignment of envelope protein sequences. Majority consensus sequences of CERV2 envelope, RhERV2 envelopes A and B, and CERV1 envelope are aligned with the moloney murine leukemia virus (moMLV) ecotropic envelope. The first and last residues of the mature surface protein (SU or gp70) and transmembrane protein (TM or p12E) are marked with red and blue asterisks, respectively. Proceeding SU is the cleaved signal peptide and following TM is the R-peptide (or p2E). Variable regions A and B are indicated by brackets according to those defined for MLV envelopes.

Figure 29

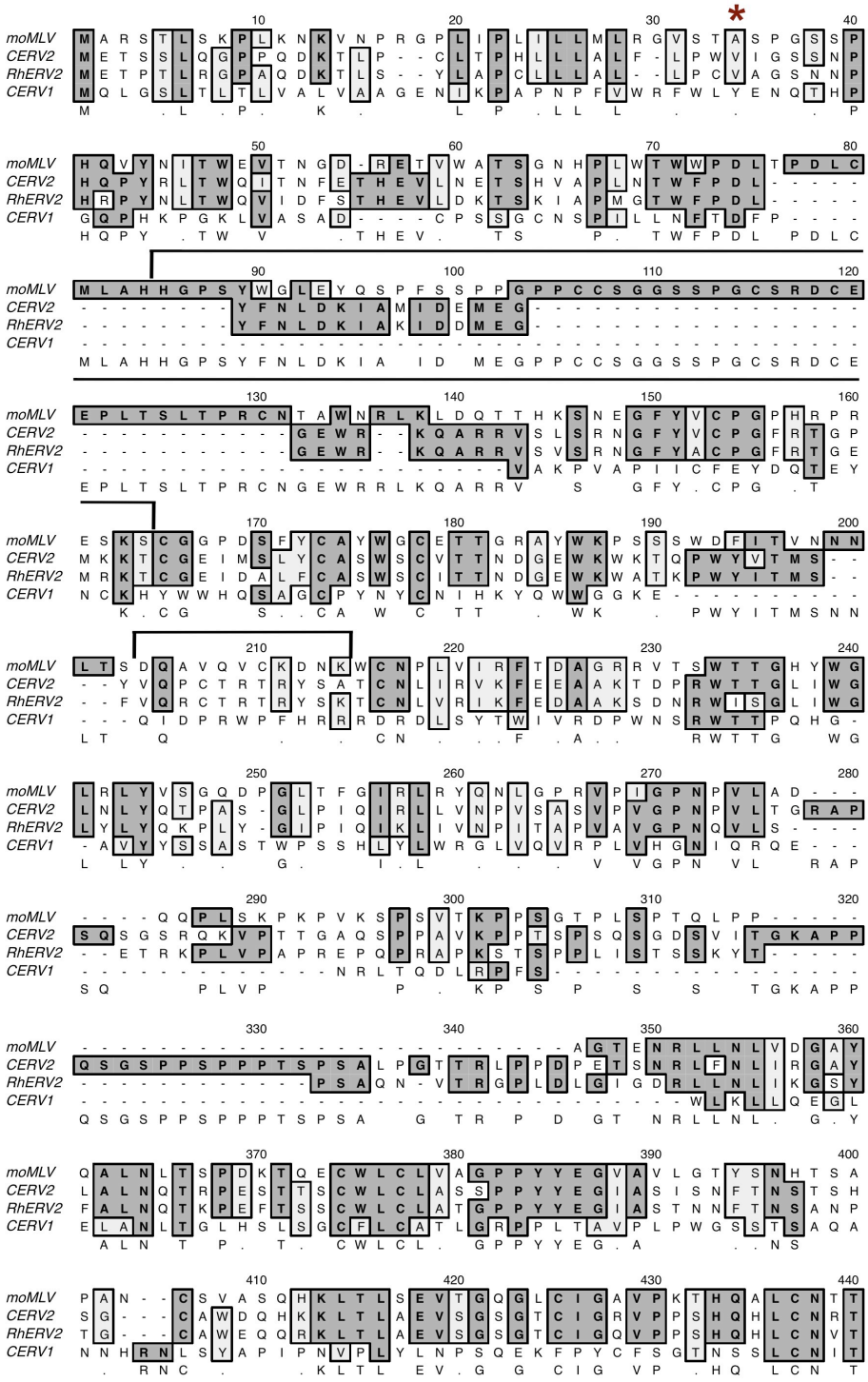
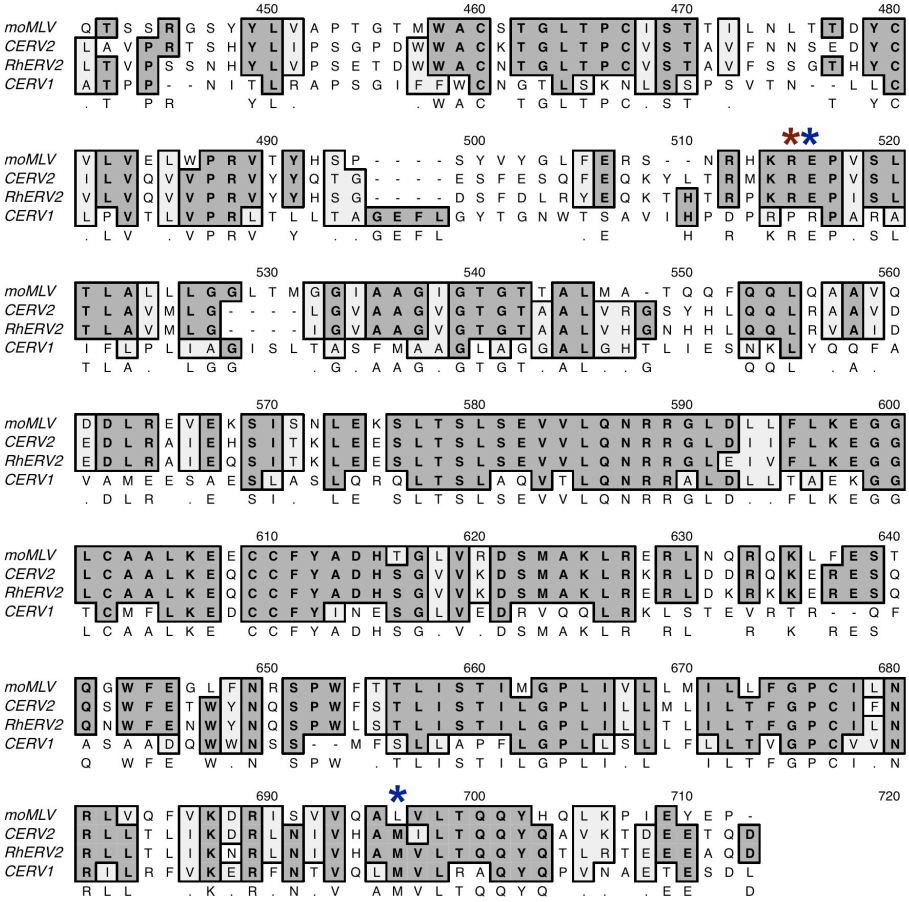


Figure 29 continued



Using CERV2 enveloped MLV pseudotypes carrying a GFP reporter (MLV-GFP), it was determined that the consensus CERV2 envelope was functional (Figure 30A). Furthermore, it displayed a broad species tropism in that it supported infection of human, chimpanzee, rhesus macaque, African green monkey, owl monkey, dog, rat, and mouse cell lines. The permissivity of HeLa cells for CERV2 infection validated the use of a HeLa cDNA library in a genetic screen for a CERV2 receptor, while the non-permissive CHO-PGSA cells provided a negative background in which to perform such a screen. CERV1 enveloped MLV particles were also able to infect both human and rhesus macaque cells, but their low infectious titer prevented any practical genetic screen for a CERV1 receptor (Figure 30C). Importantly, the ability of CERV1 and CERV2 enveloped particles to infect human cells already argues against the notion that the absence of functional receptors protected the human lineage from CERV1 and CERV2 infection.

In order to identify the receptor used by CERV2 during the time of its exogenous replication, a HeLa cDNA cell expression library was screened using particles that carried the resurrected CERV2 envelope. First, VSV-G enveloped MLV particles were used to stably transduce CHO-PGSA cells, which are normally resistant to CERV2 infection, with the HeLa cDNA library. The cDNA was delivered with an MLV vector and 3' to the cloning site, an internal ribosomal entry site (IRES) drove the expression of a zeocin resistance gene. Transduction with the cDNA library was done in 20 parallel pools of $2X10^5$ cells at an MOI of 0.5. Following zeocin selection, 10^5 cells from each PGSA cell expression library were challenged in duplicate with CERV2-enveloped virions carrying a neomycin-resistant gene. This was done at a maximal MOI (3.5) to

insure redundant coverage of every transduced PGSA clone. All but one of the forty independent infections yielded neomycin-resistant colonies ranging in number from one to twenty-four (Table 2). The colonies were pooled into 20 separate groups according to the 20 initial cell expression sub-libraries. Each of these neomycin-resistant pools was challenged independently with CERV2-enveloped virions carrying a hygromycin-resistant gene. Numerous hygromycin resistant cells readily grew out of three of the twenty pools. When challenged with CERV2-enveloped MLV virions that carried a DsRED reporter gene, the hygromycin-resistant cells displayed a 100-fold increase in permissivity relative to unmanipulated CHO-PGSA cells. Furthermore, this increase in CERV2 permissivity occurred without a change in resistance to amphotropic MLV infection (Figure 30B).

PCR amplification was performed using genomic DNA from the three PGSA cell lines permissive to CERV2 infection (Figure 30D). To amplify the transduced gene(s), primers were designed to bind sequences flanking the cloning site in the vector used to deliver the HeLa cDNA library. A 1.7kb PCR product, which was not amplified when DNA from unmanipulated PGSA was used, was cloned from all three cell lines and identified by sequence homology to encode human copper transport protein 1 (CTR1) (296). By comparison to database cDNA sequence, it was clear that all three clones were missing the same number nucleotides from the end of the 5' untranslated region, implying that the three copies of *Ctr1* isolated independently in the screen represented a single molecular clone in the cDNA library that was transduced into PGSA cells three independent times, each in a separate pool of starting PGSA cells.

Table 2. Number of neomycin-resistant colonies from CERV2-challenged CHO-PGSA cell expression libraries.

Pool	Infection A	Infection B	Total colonies
1	0	1	1
2	20	23	43
3	6	11	17
4	11	8	19
5	24	7	31
6	8	3	11
7	4	5	9
8	6	3	9
9	4	5	9
10	6	16	22
11	2	19	21
12	17	6	23
13	8	1	9
14	2	7	9
15	11	2	13
16	5	6	11
17	1	2	3
18	4	12	16
19	22	7	29
20	9	10	19
		Average:	16

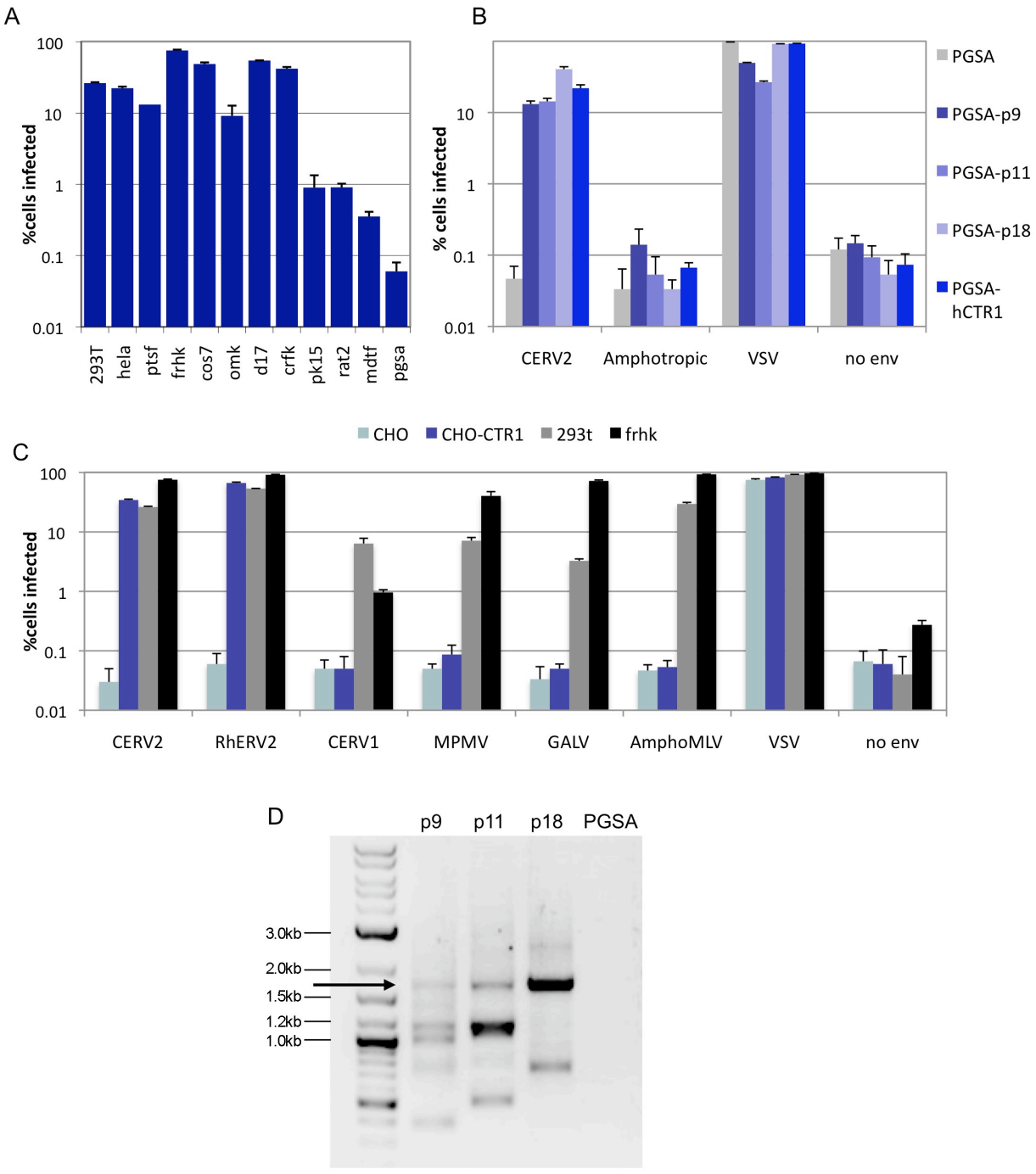
Neomycin selection was applied following challenge with CERV2 enveloped MLV virions. Results from two replicate infections of each pool are shown.

In order to confirm that CTR1 expression was sufficient to confer CERV2 sensitivity to CHO-PGSA cells, the coding region was stably transduced into both PGSA and the parental CHO cell line. The *Ctr1* open reading frame was amplified using a plasmid containing *Ctr1* cDNA and the product was cloned into an MLV vector. VSV-G enveloped MLV particles were used to transduce CHO and CHO-PGSA with both the CTR1 open reading frame and a neomycin resistance gene. Indeed, expression of CTR1

in both CHO and CHO-PGSA cells resulted in permissivity to CERV2 enveloped MLV (Figures 30B,C). Expression of human CTR1 also made CHO cells permissive to an MLV pseudotype carrying a consensus envelope from the CERV2 homologue in the rhesus macaque genome (RhERV2). In contrast, transduction of human CTR1 did make CHO cells permissive to CERV1, MPMV, GALV, or amphotropic MLV enveloped virions, providing evidence that human CTR1 specifically enhances sensitivity to CERV2.

Figure 30. Human CTR-1 is sufficient to confer permissivity to CERV2 env mediated entry in hamster cells. A) 50μL of CERV2-enveloped MLV particles carrying a GFP reporter were used to infect 10^4 cells. Cell lines from a variety of species were used- human (293T, Hela), chimpanzee (PTSF), rhesus macaque (FRHK), African green monkey (COS7), owl monkey (OMK), dog (D17), cat (CRFK), pig (PK15), rat (Rat2), mouse (MDTF), and hamster (PGSA). Cells were analyzed by FACS two days after infection. B) Following two rounds of selection with CERV2 enveloped MLV pseudotypes, cells from three separate pools of cDNA library transduced CHO-PGSA were tested for CERV2 permissivity compared to un-manipulated cells. MLV particles carrying the indicated envelope and DsRED reporter construct were used to infect cells as in A. CHO-PGSA stably transduced with human CTR1 were included in subsequent replicates. C) CHO, CHO stably transduced with human CTR1, 293T, and FRHK were infected as in A and B with MLV-GFP reporter virus pseudotyped with the indicated envelope. D) DNA from three different CHO-PGSA cell lines, shown in (B) to be permissive to CERV2, or unmanipulated CHO-PGSA were used in PCRs with primers designed to amplify transduced HeLa cDNAs. PCR products were resolved on an agarose gel. A 1.7kb product (indicated with an arrow) was common to PCRs from all cell lines except unmanipulated CHO-PGSA.

Figure 30


A
100
10
1
0.1
0.01
%cells infected
293T
hela
ptsf
frhk
cos7
omk
d17
crfk
pk15
rat2
mdtf
pgsa
B
% cells infected
CERV2
Amphotropic
VSV
no env
PGSA
PGSA-p9
PGSA-p11
PGSA-p18
PGSA-hCTR1
C
CHO
CHO-CTR1
293t
frhk
%cells infected
CERV2
RhERV2
CERV1
MPMV
GALV
AmphoMLV
VSV
no env
D
p9
p11
p18
PGSA
3.0kb
2.0kb
1.5kb
1.2kb
1.0kb

CERV2 entry into primate cells requires CTR1 and occurs at the plasma membrane

Next, we asked if CTR1 is necessary in human cells for CERV2 infection. Two different human cells lines were transfected twice with an siRNA pool targeting the human Ctr1 open reading frame or luciferase siRNA and then re-plated for infection. CTR1 siRNA pool was validated independently by knockdown of tagged CTR1 ectopically expressed in 293T cells (Figure 31A). Transfection of siRNAs against CTR1 reduced CERV2 infection by up to three-fold compared to cells transfected with control siRNA, while susceptibility to amphotropic or VSV-G enveloped MLV was unaffected (Figure 31B). To further demonstrate the necessity of CTR1 in CERV2 infections, we tested whether its normal ligand, copper, is capable of inhibiting its function as a viral receptor. Copper II chloride treatment did in fact cause a dose-dependent decrease in HeLa infection that was specific for CERV2 enveloped virions (Figure 31C).

It has been shown that ectopically expressed CTR1 is endocytosed in response to increased copper concentrations (297). In order to ask whether endocytosis, followed by lysosomal maturation is involved in CERV2 entry, we determined if CERV2 infection is pH-dependent. Pre-treatment and infection in the presence of either ammonium chloride or the lysosomal maturation inhibitor, bafilomycin-A demonstrated in CHO-hCTR1 and TE671 that infection by CERV2 enveloped MLV is a pH-independent event (Figure 32).

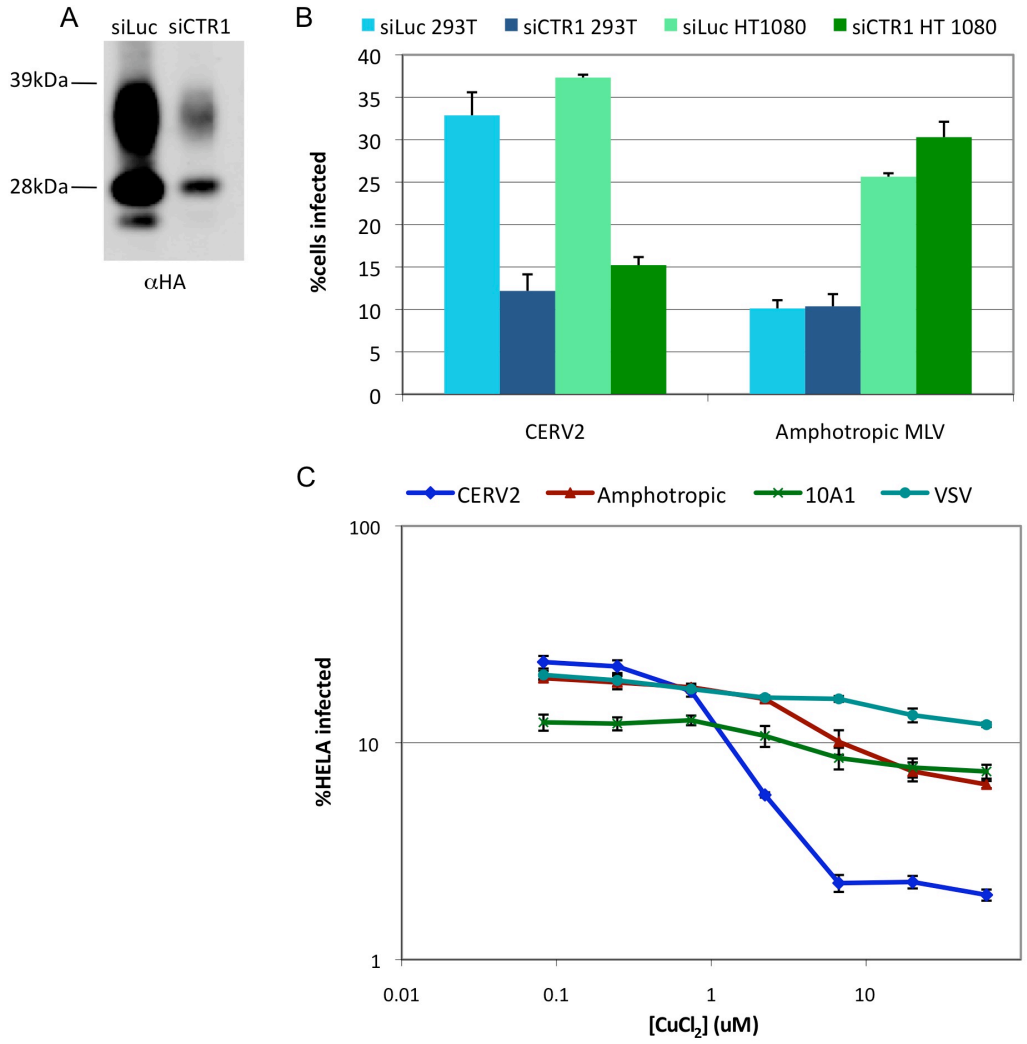


Figure 31. CTR1 is necessary for CERV2 infection in human cells. A and B) 293T or HT1080 cells were transfected twice with an siRNA pool targeting CTR1 or luciferase siRNA before seeding for infection with MLV-GFP carrying CERV2 or amphotropic envelope. Separately, 293T cells were co-transfected with plasmid expressing human CTR1 with an N-terminal HA tag. Western blots from these cells are shown in A. Glycosylated forms of CTR1 can be observed above the un-glycosylated protein (298-300). C) HeLa cells were incubated for 2 hours in $CuCl_2$ before infection at the same $CuCl_2$ concentration with the indicated MLV-GFP pseudotype. After overnight incubation, cells were washed with fresh media and FACs analysis was completed 2 days after infection.

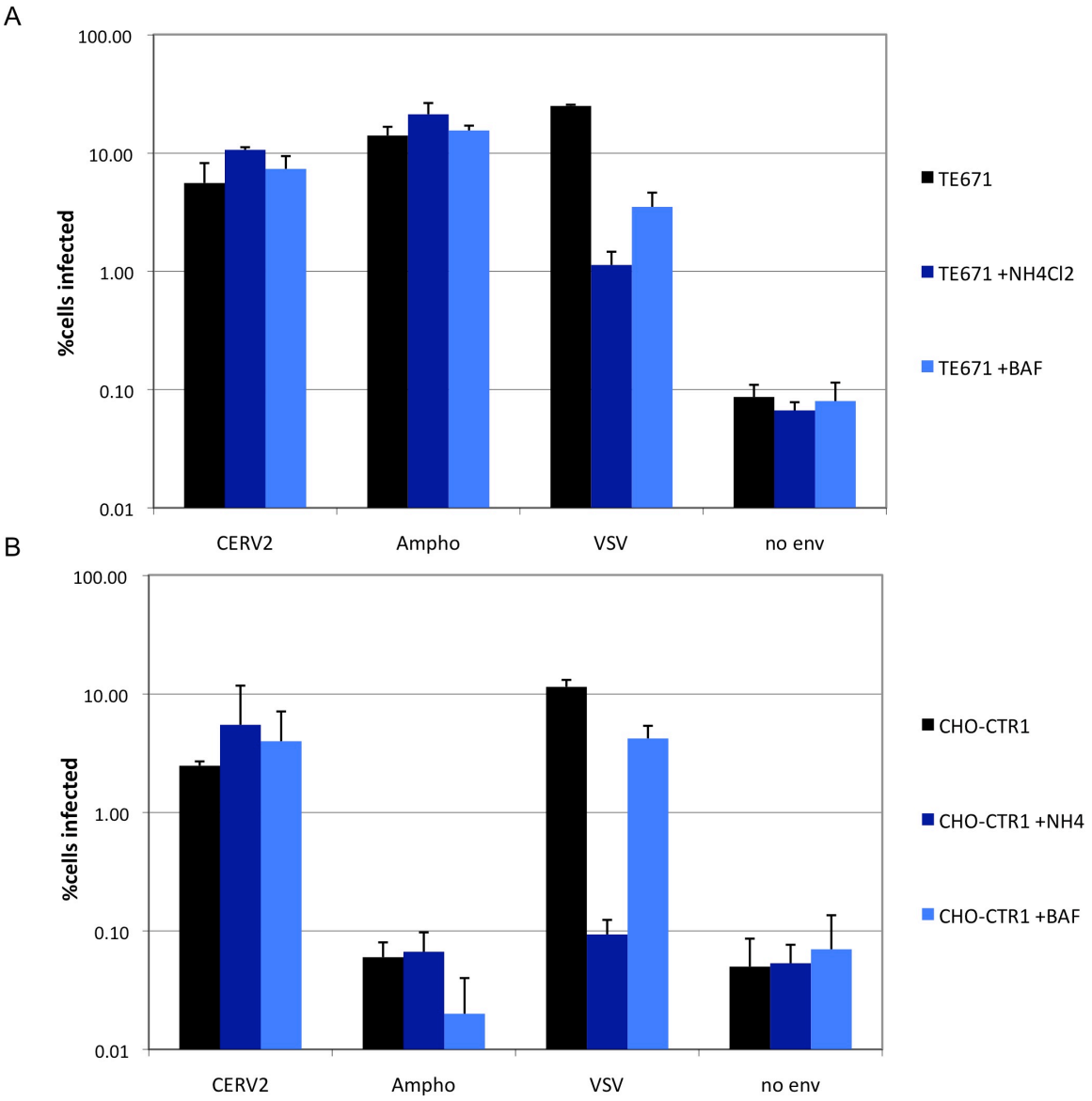
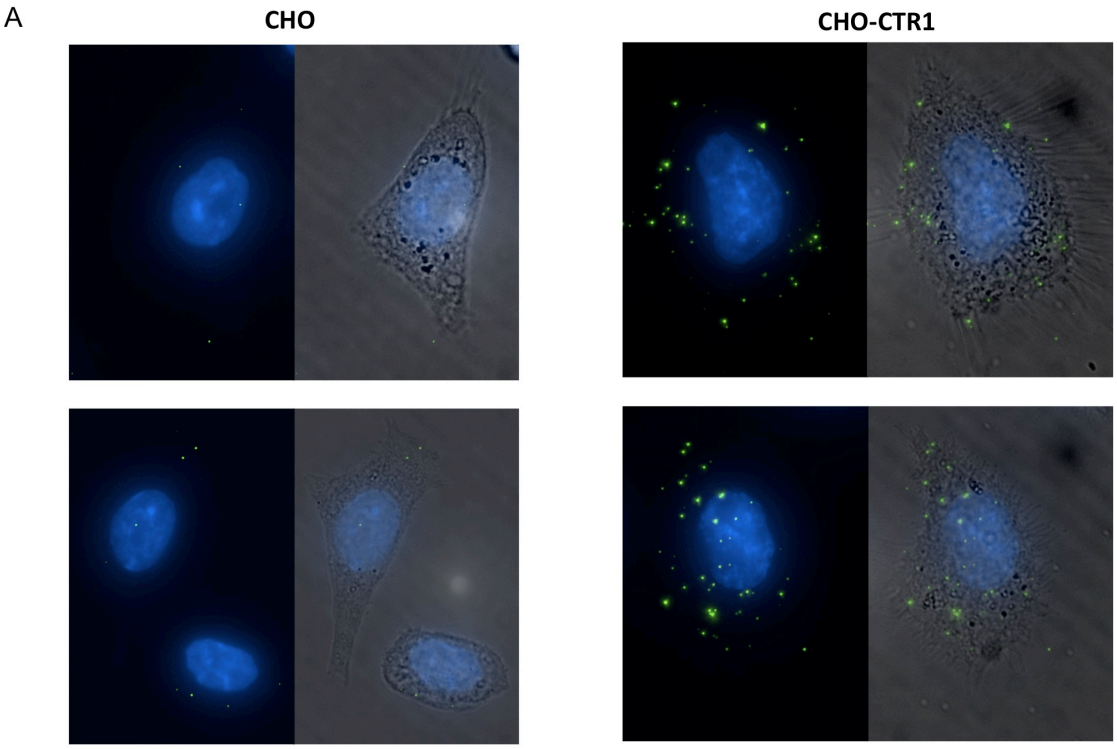
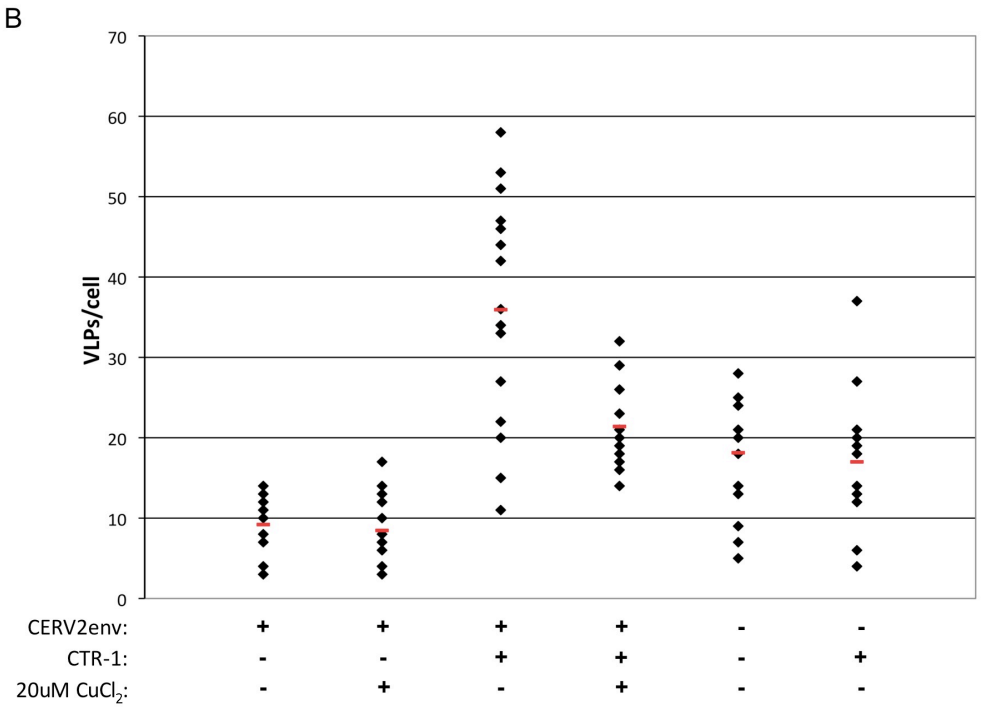


Figure 32. CERV2 entry is pH-independent. TE671 (A) or CHO-huCTR1 (B) cells were infected by CERV2, amphotropic, or VSV enveloped MLV-GFP +/- 30mM NH_4Cl or 25nM bafilomycin A following a two-hour pre-incubation under the same condition. After 3 hours of infection, virus was removed and cells were incubated for an additional 2 hours in fresh media with 30mM NH_4Cl or 25nM bafilomycin A before replacement with plain media.

Figure 33. Human CTR1 promotes binding of CERV2 particles to CHO cells. MLV Gag-GFP particles with or without CERV2 envelope were bound to CHO or a CHO-humanCTR1 single-cell clone. After one hour, cells were washed, fixed, and stained with DAPI for microscopy. In parallel, CERV2 particle binding was also done in the presence of $CuCl_2$ following a $CuCl_2$ pre-incubation. Images of CERV2 particles bound to each of the cell types are shown including a merge of fluorescence with phase contrast. For quantitation, 15 cells from each condition were chosen at random and the number of GFP particles per cell was counted. The mean is indicated by a bar.

Figure 33


A
CHO
CHO-CTR1


B
VLPs/cell
70
60
50
40
30
20
10
0
CERV2env: + + + + - -
CTR-1: - - + + - +
20uM CuCl2: - + - + - -

CTR1 induces binding of CERV2 to the plasma membrane

To visualize virus attachment to target cells, an MLV Gag was used that carried a C-terminal GFP fusion. Plasmid encoding the MLV Gag-GFP was transfected either alone or with plasmid expressing the CERV2 envelope. This resulted in the production of GFP-labeled virions that either carried CERV2 envelope or lacked an envelope glycoprotein. Particles were bound to CHO cells or a CHO single-cell clone expressing human CTR1 by applying un-concentrated supernatent at 4°C. Following a one-hour incubation with the labeled virions, cells were washed three times and then fixed for microscopy. Binding of CERV2 enveloped virions to unmanipulated CHO cells, in terms of the average number of GFP foci per cell, was seen only at a very low level- lower even than non-specific binding by particles lacking envelope glycoprotein (Figure 33B). On the other hand, numerous CERV2 enveloped particles were bound to CHO-CTR1 cells (Figures 33A, B). There were four times more CERV2 particles bound to CHO-CTR1 compared to CHO, and expression of human CTR1 did not have an effect on the binding of non-enveloped virions. This demonstrates that human CTR1 is able to confer CHO cells with the ability to bind CERV2 virions.

To further demonstrate that CTR1 is the cellular receptor for CERV2, we asked if excessive copper decreases virion binding to cells (Figure 33B). CHO or CHO-CTR1 cells were pre-treated with 20μM $CuCl_2$ followed by binding with CERV2-enveloped GFP-labeled virions in the presence of 20uM $CuCl_2$. Treatment with $CuCl_2$ caused an approximately two-fold decrease in the average number of CERV2 virions bound to CHO-CTR1. This effect was modest compared to the 10-fold effect that 20uM $CuCl_2$ had

on HeLa infection. Therefore, it is likely that the binding defect does not entirely explain the effect that copper treatment has on CERV2 infection and, perhaps, excessive copper can also perturb events downstream of CTR1 binding.

Genetic determinants of CTR1 receptor function

To determine if the use of CTR1 by CERV2 is directly related to its cellular function, mutants that have been previously shown to be deficient in copper transport were tested for viral receptor function. Two CTR1 mutants with single amino acid changes that were previously shown to have an effect on the rate of copper uptake were included: M81I and Y156A (301). A third included point mutation, C189S, has been shown to decrease CTR1 trimer stability (298). None of these point mutations affected the ability of human CTR1 to serve as a CERV2 receptor (Figure 34B). Therefore, the ability of CTR1 to serve as a receptor for CERV2 can be uncoupled from its copper transport function.

Next, we sought to identify the genetic determinants of CTR1 viral receptor function that render hamster cells non-permissive, and human cells permissive to CERV2 infection. Stable transduction of CHO cells with hamster CTR1 did not confer CERV2 permissivity, suggesting that the lack of CERV2 infection in unmanipulated CHO cells was not merely due to low endogenous CTR1 expression (Figure 34A). CTR1 has three transmembrane domains with a 64 amino acid extracellular N-terminal domain and a three amino acid extracellular loop (298, 299, 302). To investigate whether the N-terminal extracellular domain from human CTR1 is both sufficient to confer viral receptor activity to hamster CTR1 and necessary for the viral receptor function of human

CTR1, this domain was swapped between the human and hamster genes and the resulting chimeras were transduced into CHO cells. CHO cells stably expressing a hamster CTR1 with the human N-terminal extracellular domain were as permissive to CERV2 as CHO cells expressing the full-length human CTR1, whereas CHO cells expressing human CTR1 with the hamster N-terminal extracellular domain were non-permissive to CERV2 (Figure 34A). A C-terminal HA tag was included with the human, hamster, and chimeric CTR1 proteins and a western blot demonstrated that all four cell lines had comparable levels of CTR1 expression (Figure 34A insert).

Alignment of CTR1 amino acid sequences from primates revealed a high degree of conservation (Figure 35). CTR1 proteins from human and chimpanzee are identical and, relative to this amino acid sequence, the rhesus macaque CTR1 has two amino acid differences and the orangutan, only one. All three of these mutations are found within the N-terminal area of the extracellular domain that precedes an essential MMMMPM copper coordination motif (303). An alignment of CTR1 from a wide range of mammals, frog, and zebrafish shows broad divergence in this 40 amino acid N-terminal region, which is followed by a highly conserved sequence in the rest of the protein. Transduction of CHO cells with either wildtype human CTR1 or a mutant with a 40 amino acid N-terminal truncation was able to induce CERV2 permissivity to a similar degree (Figure 34B). The ability of CERV2 to use the CTR1 truncation mutant as a receptor demonstrated that the most divergent region of CTR1 is dispensable in its receptor function. As was therefore expected, hamster CTR1 mutants that resembled human CTR1 in the membrane proximal portion of the extracellular domain were competent as CERV2 receptors. Specifically, stable expression of a hamster CTR1 that contained the

R(53)N and Y(59)S mutations (numbered by homology to human CTR1) was able to induce CERV2 susceptibility in CHO cells. Furthermore, the addition of three methionines to form a MMMMXM motif improved the receptor function of the R(53)N/Y(59)S hamster CTR1 mutant. Inversely, a three-methionine deletion in the human CTR1 MMMMXM motif resulted in decreased receptor activity and the N53R/S59Y human CTR1 double mutant did not display any receptor function in CHO cells. An N-terminal HA tag was included in all CTR1 proteins tested and expression in transduced CHO cells was confirmed by western blot and surface stains (Figure 34B insert and Table 3). Both the human CTR1 mutant containing a three-methionine deletion and the N53R/S59Y double mutant had lower expression levels in CHO cells compared to the wildtype human CTR1. It is possible that these expression differences partially explain their lack of receptor activity. Nevertheless, the gain of receptor function induced by the reciprocal mutations in hamster CTR1 clearly demonstrates the importance of these residues in CERV2 infection.

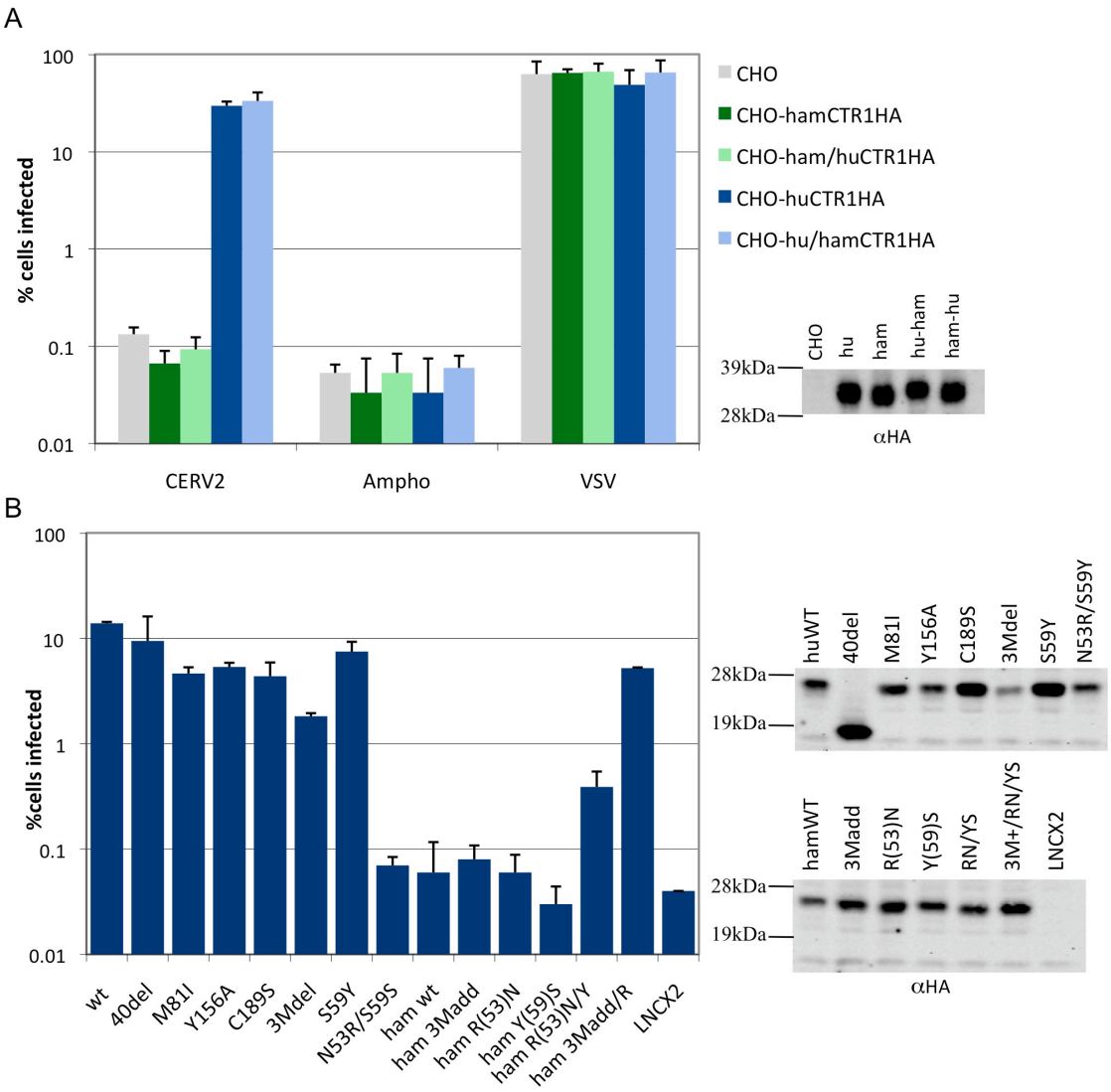


Figure 34. Genetic determinants of CTR1 viral receptor function. A) CHO cell lines expressing C-terminal HA tagged human CTR1 (Hu), hamster CTR1 (Ham) or N-terminal extracellular domain swaps (Hu-Ham and Ham-Hu) were seeded for infection with MLV-GFP pseudotypes. Expression of CTR1 was detected by αHA western blot, and the glycosylated band is shown to the right. B) CHO cells expressing N-terminal HA tagged CTR1 were infected with CERV2 enveloped MLV-GFP or used for αHA western blot (shown on right). Mutations to human CTR1 (wt) or hamster CTR1 (hamster wt) are indicated including a deletion of 40 amino acids at the N-terminus (40del) of human CTR1, deletion of methionine residues 41-43 (3Mdel) in human CTR1, or addition of three methionine residues at the orthologous position in hamster CTR1 (ham 3Madd).

Table 3. Surface staining of CHO cells expressing N-terminal HA tagged CTR1.

CHO cell lines	MFI	%cells +
wt human CTR1	161	62.5
40del	154	84.4
M81I	324	94.4
Y156A	107	41.2
C189S	339	27.9
3Mdel	109	37.5
S59Y	342	79.0
N53R/S59S	92	40.0
wt hamster CTR1	160	73.9
ham 3Madd	92	45.8
ham R(53)N	135	58.3
ham Y(59)S	153	63.3
ham R(53)N/Y(59)S	125	52.4
ham 3Madd/R(53)N/Y(59)S	203	78.0
LNCX2	9	1.2
wt human CTR1 unstained	3	0.5

Stains were done with a mouse αHA monoclonal antibody and an AlexaFluor488 conjugated secondary antibody before FACS analysis. Mean fluorescence intensity (MFI) of the entire population and the %cells with fluorescence greater than that of vector and unstained controls is shown.

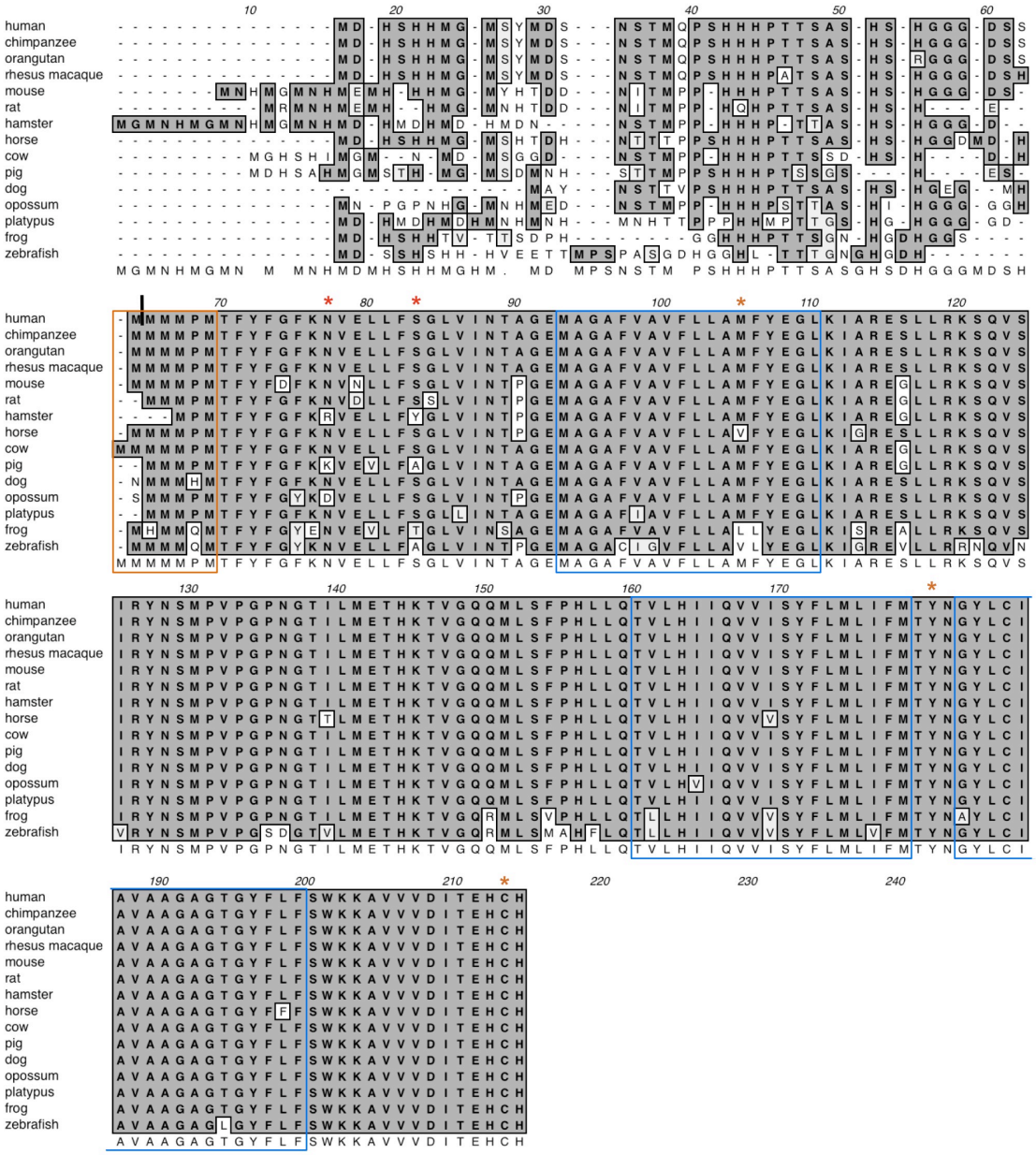


Figure 35. Clustal alignment of CTR1 proteins. The MMMMXM copper-coordination motif is marked by a copper colored box, transmembrane domains by blue boxes, and residues mutated for the studies herein are indicated by asterisks. The hamster CTR1 sequence was determined by cloning the gene from CHO cDNA. All other sequences were previously annotated in the NCBI gene database or were collected by a TBLASTN search of the non-redundant nucleotide database. Sequences from chimpanzee, rhesus macaque, horse, dog, platypus, and opossum are annotated in the sequence database as predicted transcripts. In the case of dog CTR1, a start codon at a position homologous to other species is lacking therefore a downstream hypothetical start codon (proceeding a G nucleotide) is shown.

Chapter VI. Discussion

The study of contemporary viruses can describe modern, but not necessarily formative, interactions between host genes and pathogens. The existence of endogenous retroviral relics in host genomes allows for explorations of host and retrovirus co-evolution that are not limited to present day infectious agents. The work herein endeavored to gather evidence for or against past encounters of host factors with viruses that have not replicated exogenously for millions of years, but may have imposed positive selection on primate genes. The peculiar absence of CERV1 and CERV2 from the human genome made them intriguing subjects for such studies because restrictions to their exogenous replication by antiviral proteins may have limited their host ranges or even caused their extinction.

Our investigation began with TRIM5 and APOBEC3 proteins because their ability to restrict modern MLV set a reasonable precedent for the hypothesis that they were also capable of targeting archaic primate gammaretroviruses (109, 110, 160, 304). Evidence argued against the role of TRIM5α proteins, though, in both the inactivation of CERV1 and CERV2 and in broad restrictions to their cross-species transmissions. Nearly all endogenous capsid N-terminal domains studied lacked sensitivity to TRIM5 proteins from a variety of primate species. RhERV2 capsids carrying an E91K mutation were an exception to the general TRIM5 resistance displayed by all other endogenous retroviral capsids. Virions carrying these particular RhERV2 capsids were restricted by chimpanzee TRIM5α and marginally by human TRIM5α. The abundance of 91K in

RhERV2 capsids (~20%) and its complete absence in CERV2 suggests that TRIM5α may have imposed a selective pressure that limited the cross-species transmission of viruses carrying this particular mutation to the chimpanzee ancestor and/or imposed a constraint on CERV2 capsid variability.

There are two possible caveats to this study. First, the inactivation of CERV1/2 viruses, and hence the end of any possible selective pressure imposed by these retroviruses on TRIM5 genes, may have occurred early enough in evolutionary history to allow genetic drift and/or further selection by more recent pathogens to cause a loss in recognition of ancient capsids by TRIM5 proteins. In other words, neutral evolution or positive selection imposed by viruses other than CERV1/2 may have caused TRIM5 proteins to lose CERV1/2 capsid recognition. Contrary to this argument, though, is the high degree of functional conservation and sequence identity between human and chimpanzee TRIM5α: 97% by amino acid sequence. Since the replication of CERV1/2 occurred after the divergence of humans and chimpanzees six million years ago, any recent changes in TRIM5 specificity caused by the rise and fall of species-specific retroviruses should be reflected in the divergence of these two genes; but despite the absence of CERV1/2 from the human genome and its presence in the chimpanzee, TRIM5α remains heavily conserved in both sequence and functional terms in these two species. Furthermore, there are polymorphisms that remain common to both the rhesus macaque and the sooty mangabey despite geographic separation and eight million years of divergence, which suggests that balancing selection has prevented primate TRIM5α from significantly shifting during the last few million years (305). In short, it is likely that the replication of CERV1/2 and their homologues is too recent relative TRIM5

functional divergence for changes in TRIM5α specificity to pose a problem in the experiments discussed here.

Second, it is possible that there are unknown determinants of TRIM sensitivity that are outside of the N-terminal domain of capsid. This concern was especially highlighted by the description of a chimeric MLV containing p12 and a complete capsid protein of an ancestral CERV1 that was restricted by human and chimpanzee TRIM5α (260). It seems unlikely that the inclusion of the CERV1 CA-CTD or the p12 domain could affect TRIM5 sensitivity because all known determinants of TRIM5 and Fv1 sensitivity map to the CA-NTD (108-111, 113, 114), and cleavage from the p12 domain is required for MLV CA to interact with TRIM5 and Fv1 (306). However, in order to deal with the possibility that there is an unknown determinant of TRIM5 recognition in the capsid C-terminal domain, ancestral and consensus CERV1 full-length capsids were used to confirm that chimeras carrying CERV1 capsid are resistant to TRIM5. Thus, this result strongly supports the notion that TRIM5α proteins did not restrict CERV1 and it also validates our experimental strategy for testing a broader range of CA-NTDs for TRIM5 sensitivity.

Although it appears that TRIM5 proteins did not have an impact on CERV1 and CERV2 replication, evidence does suggest that these viruses encountered APOBEC3 family member in both the chimpanzee and the rhesus macaque. This is evidenced by a large number of G to A changes that occurred in dinucleotide contexts that are indicative of modern APOBEC3 proteins. A disproportionate number of GG to AG mutations compared to other G to A mutations was documented in CERV1, CERV2, RhERV1a, RhERV2 and an excess of GA to AA mutations were also observed in individual

RhERV2 proviruses. It is intriguing that RhERV2, unlike the other groups studied, displayed two distinct dinucleotide biases: both GG to AG and GA to AA mutations on separate proviruses. It is unlikely that these two mutagenic activities were separated temporally because the proviruses in which they were found are not separated into two phylogentically distinct groups. Rather, it is likely that RhERV2 was simultaneously subjected to the action of at least two different deaminases. Indeed, modern rhesus APOBEC3F and APOBEC3H are highly active relative to their human orthologues, suggesting that they may be responsible for the abundant GA to AA mutations found in RhERV2 (155-158).

Several analyses were completed in order to confirm that the excessive G to A mutations in GG or GA dinucleotides were caused by APOBEC3-catalyzed deamination. By removing CG dinucleotides from our analyses, we confirmed that the GG to AG mutations observed were not due to spontaneous deamination at methylated CG islands in the context of CGG trinucleotides. Some of the G to A mutations that were common to multiple proviruses may have been due to amplification of a relatively small number of G to A changes from founder viruses. It is unlikely, though, that the GG to AG changes would be selectively amplified over other mutations because these changes did not confer any fitness advantage MLV chimeras containing CERV1/2 CA-NTDs. To confirm that the dinucleotide biases observed extend outside of capsid, many of the results were corroborated by examining mutational biases in envelopes that were linked to the capsids analyzed. Cases in which the bias in the G to A mutations was not consistent between a capsid and a linked envelope could be explained by (i) stochastic variation in the small numbers of capsid G to A mutations in th relatively short CA-NTD sequence (~400nt) (ii)

variation in the mutation frequency (including 5' to 3' gradients) across individual proviruses (iii) recombination between hypermutated and intact proviruses prior to or during germline invasion, and/or (iv) high degrees of general sequence diversity in some viral genes that may mask low frequency APOBEC3-mediated editing.

Modern APOBEC3G is known to induce GG to AG mutations but it remained possible that an un-characterized cytidine deaminase activity was responsible for these mutagenic events in CERV1 and CERV2. Furthermore, direct evidence that the editing observed in these viruses translated to an observable restriction was lacking. To address this, it was demonstrated *in vitro* that APOBEC3G can, in fact, restrict MLV chimeras carrying CERV2 Gag. This supports the assertion that APOBEC3G was responsible for GG to AG mutation found in CERV2 and that this activity was capable of restricting CERV2 replication.

APOBEC3-induced mutations often crippled individual CERV1 and CERV2 viruses through the frequent introduction of stop codons and other deleterious nonsynonomous changes. Especially given high frequency of stop codons generated by GG to AG mutations, it is possible that APOBEC3 activity resulted in genetic loads that had a damaging impact on the viral populations within each infected individual and therefore may have also decreased the likelihood of intra- and interspecies transmissions. Especially given the *in vitro* results confirming that APOBEC3G is able to restrict the replication of virions containing CERV2 Gag , it is possible that APOBEC3 genes were able to limit the species tropism of CERV1 and CERV2. Indeed, the species-specificity of APOBEC3 recognition by lentiviral Vif proteins at least partially explains the species tropism of modern lentiviruses; thus demonstrating the potential for APOBEC3 proteins

to impose transmission barriers (156, 175, 307). *in vitro* APOBEC3G restriction assays, though, were unable to provide direct evidence for a species-specific difference in APOBEC3G potency against CERV2. Specifically, virions containing CERV2 Gag were not more sensitive to human APOBEC3G compared to chimpanzee APOBEC3G, which was apparently not potent enough to prevent CERV2 endogenization in the chimpanzee ancestor.

It remains possible that the APOBEC3 activity described here was responsible for terminating CERV1/2 replication. Again, deleterious non-synonymous mutations, particularly stop codon formation, likely constituted a genetic load capable of impairing the replication of a viral population within an infected host. An assertion that these genetic loads caused CERV1/2 extinction may be premature, though, given that it is not known if the level of mutations observed in CERV1/2 constitute a burden that was too great for the viruses to overcome. It has long been theorized that recombination evolved, in part, as a means to rid genomes of deleterious mutations, thus avoiding Muller's ratchet and preventing mutational meltdown (204-207). CERV1/2 may have been able to cope with the heavy mutational burden observed here by continually producing fit viruses through the recombination of mutated sequences. It should also be stated that the proportion of proviruses in the infected chimpanzee and rhesus macaque ancestors that were hypermutated would be over-represented if a selection existed against the endogenization of fit viruses. Overall, although APOBEC3 activity likely explains the inactivity of many individual CERV1/2 proviruses, it may or may not have contributed to the extinction of exogenous CERV1/2.

Further study of the CERV2 host range, as well as the limitations that hosts may have imposed on it throughout history, required the identification of its cellular receptor. By resurrecting a functional CERV2 envelope, we have determined that CERV2 used CTR1 as its cell surface receptor during its exogenous replication millions of years ago. Human CTR1 expression was sufficient to allow CERV2 virion binding and infection in otherwise resistant hamster cells. In human cells, siRNAs directed against CTR1 inhibited CERV2 infection. The remaining infection in CTR1 knockdown cells could be due to an incomplete knockdown but we cannot exclude the existence of a second CERV2 receptor. The large effect that $CuCl_2$ treatment had on CERV2 infection further supports the importance of CTR1 in the infection of human cells. Although potent inhibition by $CuCl_2$ was clearly specific for the CERV2 receptor, the exact mechanism is unclear. Down-regulation of ectopically expressed CTR1 from the cell surface in response to $CuCl_2$ treatment has been observed but it has also been shown that there is no such effect on endogenously expressed CTR1, suggesting that inhibition of HeLa cell infection is not necessarily explained by a lack of CTR1 cell surface expression (308-310). Furthermore, the modesty of the binding defect caused by $CuCl_2$ suggests that perturbation of events downstream of initial binding to CTR1 may contribute to the inhibitory effect that copper treatment had on infection. It is reasonable to speculate that high amounts of copper flux at or near sites of virion binding may interfere with proper virus attachment and/or fusion with the cell.

Several γ-retroviruses use multi-transmembrane transport molecules as cell-surface receptors. Aside from their expression on the cell surface, no common property of transport proteins has been shown to promote their ability to serve as receptors for

viruses. In the case of CTR1, we demonstrated that viral receptor activity could be uncoupled from its normal cellular function. Methionine residues are primarily responsible for copper ion coordination during uptake through the membrane pore formed by trimerized CTR1 (302, 303, 311). Mutation at methionine 81 within the first transmembrane domain has been demonstrated to result in a two-fold decrease in the Vmax of copper uptake (301). It has been suggested that tyrosine 156, found in the three-residue extracellular loop, is responsible for positioning methionine residues at the exterior opening of the transport pore through hydrogen bonding and mutation at this residue results in a five-fold loss in copper uptake (301). The observation that point mutants lacking each of these residues were able to serve as efficient viral receptors supports the view that, as has been shown for the ecotropic MLV receptor, solute transport function is not linked with virus reception (312).

The ability of CERV2 to use human CTR1 suggests that a species-specific tropism during the period of exogenous CERV2 replication does not explain the absence of CERV2 from the human genome. Additionally, human CTR1 is broadly expressed (313), which is to be expected given the use of copper ions by crucial cellular enzymes such as superoxide dismutase and cytochrome c oxidase. In particular, the expression of CTR1 in the testes and ovaries implies that a lack of endogenization cannot be explained by the absence of receptor expression in the germ line. Although we cannot rule out that human CTR1 has lost CERV2 resistance during the time since exogenous CERV2 replication, this evolutionary course is unlikely given the perfect nucleotide sequence identity between the extra-cellular regions of human and chimpanzee CTR1. Still, it

remains possible that CTR1 allelic variability once included CERV2 resistance and perhaps such polymorphisms were lost following the extinction of exogenous CERV2.

Interestingly, nearly all of the CTR1 sequence diversity is found in its N-terminal extracellular domain and a 40 amino acid truncation mutant demonstrated that the area containing nearly all of this diversity is dispensable for virus receptor function (Figures 34 and 35). That CERV2 only requires the extracellular domain's most conserved portion likely explains the permissivity of cells from diverse species. Furthermore, it is possible that binding to this conserved region has been selected by multiple cross-species transmission events that favored binding to a region in the new host that is similar to that used by the virus in the previous host. For sequence variability to dictate the portion of the receptor that was required for CERV2 infection, the virus would have to be capable of altering its receptor usage when faced with unfavorable residues on CTR1. Observed differences in receptor utilization that exist between closely related gammaretroviruses suggests that they are, indeed, capable of rapidly changing the way they interact with a single receptor in response to selective pressure imposed by the host. For example, polytropic and xenotropic MLV bind to the same XPR1 domain and a few differences in the receptor binding regions of xenotropic envelope enable it to also bind to a second domain on XPR1 (78). Also, following the transmission FeLV-A to a new host, FeLV-B or FeLV-C arise recurrently through either point mutations or recombination with endogenous retroviruses, thus resulting in a swift change to a new receptor utilization (86, 90).

Functional flexibility in the same extracellular region of CTR1 that is dispensable for viral infection is not only suggested by diversity between species, but also by

mutagenesis data. A human CTR1 mutant with a 34 amino acid N-terminal truncation has been shown to entirely retain its copper transport ability (301). It was suggested that this portion is responsible for receiving copper from extracellular copper chaperones that are yet to be discovered, but the data presented here raises the question of whether CTR1 divergence could have been due to selective pressure imposed by retroviruses. The extracellular copper-binding residues that are the most crucial to CTR1 function were previously demonstrated to be the highly conserved MMMMXM motif (303), which begins one residue before the end of a portion of CTR1 that is dispensable for CERV2 infection. In contrast to the high degree of conservation of this motif from human to zebrafish, the hamster CTR1 has only an MPM in the orthologous position and, perhaps for compensation, contains several unique copper-coordinating residues in other areas of its extracellular domain (Figure 35). We have demonstrated that the three-methionine deletion, along with point mutations at two other highly conserved residues, are responsible for the resistance of hamster cells to CERV2 infection. It is tempting to speculate that these changes represent an escape of the hamster lineage from a CERV2 relative. Whether or not this was the case, the divergence seen in the rest of the hamster CTR1 N-terminal domain may compensate for its loss in the primary copper coordination motif, thus representing an evolutionary hurdle that a primate would have had to overcome to acquire CERV2 resistance by way of receptor evolution. The work described here demonstrates that paleovirology has the potential to both identify novel host factors that were usurped by ancient pathogens, as well as indicate the limitations that hosts may have had in avoiding such exploitations by viral invaders.

If the human ancestors were unable to avoid CERV2 infection through changes in CTR1, then why were they spared from CERV2 endogenization? Again, we cannot rule out a role for APOBEC3-mediated restriction in the protection of the human lineage from both CERV1 and CERV2. But given the absence of clear species-specific differences in CERV2 restriction potency, such a conclusion would fail to explain why APOBEC3 was not also cable of protecting non-human primate ancestors from infection. One would have to invoke a model in which resistance to CERV2 infection was evolved in ancient primates through selection on APOBEC3 genes by CERV2 and the inactive proviruses in the chimpanzee and rhesus macaque ancestors are the products of restrictions that occurred following APOBEC3 adaptation. If replication competent CERV2 was pathogenic, then a record of proviruses that integrated prior to the emergence of APOBEC3-mediated hypermutation would be limited by selection against the endogenization of active proviruses.

The above remarks must be stated with circumspection, though, given that the presence of hypermutated proviruses does not necessarily predict that a given viral population would be unable to replicate. If we may draw analogies between the effects that APOBEC3 can have on distant retroviruses, the interactions that APOBEC3 proteins have with HIV-1 suggest that a virus capable of overcoming restriction through the use of an antagonist can still be subjected to hypermutation. Although an accumulation of APOBEC3-mediated mutations can be observed in a portion of HIV-1 populations *in vivo* and *in vitro*, the virus is still able to persist and avoid a mutational meltdown. In the *in vitro* experiments presented here, hypermutation was observed in a vast majority of sequences collected from deltaVif virus grown in restrictive cells. Importantly, the virus

was able to persist in spite of the APOBEC3-induced genetic load that likely decreased fitness and was not able to contribute to drug escape. In other words, heavy hypermutation can impair viral evolution, but it does not necessarily provide a general protection for the host or cause a mutational meltdown. Furthermore, in the case of Vif+ HIV-1, the presence of a low level of APOBEC3 activity did not appear to hinder drug resistance evolution and may have even contributed to it.

Although HIV-1 and CERV1/2 are very distant viruses, it may be a recurrent theme in APOBEC3 activity that hypermutation can occur even during the replication of a virus that is able to persist. If this was the case for CERV1/2, then the endogenized hypermutants may represent a small portion of the proviruses that has been biased toward hypermutants by selection against the fixation of replication competent individuals. Given that endogenization is a rare event and may exclude a highly fit majority, perhaps the absence of CERV1/2 in the modern human genome does not necessarily suggest that human ancestors were spared from exogenous CERV1/2. It may even be the case that polymorphic CERV1/2 integrations were lost due to population bottlenecks during human evolution. This has been suggested for a single HERV-K integration that is identical in the gorilla and chimpanzee, while the orthologous locus in the human genome contains an intact pre-integration site (314). It was argued that the integration occurred prior to the divergence of the gorilla lineage from the human and chimpanzee and remained polymorphic during the divergence of the human and chimpanzee lineages. The allele carrying the integrated provirus then became fixed in the chimpanzee lineage but, by chance, was lost from the human lineage as the allele containing the unperturbed pre-integration site became fixed. The notion that endogenous retroviruses can remain

polymorphic for long periods of time is supported by the observation that two HERV-K insertions, HERV-K113 and 115, remain polymorphic in humans (315). If CERV1/2 did infect human ancestors and endogenizations only occurred in a small proportion of individuals, than one could imagine the neutral loss of entire groups of polymorphic endogenous retroviruses to population bottlenecks.

Nevertheless, a definitive molecular explanation for the absence of CERV1/2 from the human genome may still exist. The CERV1 consensus envelope was not sufficiently functional to perform a genetic screen, but if an improved envelope were to be constructed then receptor identification would be worthwhile and could potentially account for a limited CERV1 host range. Also, the tetherin restriction factor should be considered. Although tetherin activity, unlike TRIM5α, does not appear to depend on the specific recognition of viral proteins, perhaps an unknown tetherin antagonist was able to protect primate gammaretroviruses in a species-specific manner. Another potential avenue of study ignored here is CERV1/2 LTR promoter function. Perhaps the CERV1/2 LTRs were unable to promote transcription in human ancestors. ZFP809 has been identified as a murine zinc finger protein responsible for silencing the transcription of MLV in embryonic stem cells through the bridging of transcriptional repressors to the retroviral primer-binding site (316). Perhaps similar activities have guarded the human lineage from gammaretroviruses. In general, as host genes continue to be identified that either aid or inhibit retroviruses, the history of retroviral evolution may be further explored and it is clear that paleovirology will continue to describe the formative and longstanding evolutionary combat between retroviruses and their hosts.

References

1. Temin HM (1964) Homology between Rna from Rous Sarcoma Virous and DNA from Rous Sarcoma Virus-Infected Cells. *Proc Natl Acad Sci U S A* 52:323-329.

2. Chun TW*, et al.* (1997) Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* 387(6629):183-188.

3. Siliciano JD*, et al.* (2003) Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med* 9(6):727-728.

4. Finzi D*, et al.* (1999) Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med* 5(5):512-517.

5. Weiss RA & Payne LN (1971) The heritable nature of the factor in chicken cells which acts as a helper virus for Rous sarcoma virus. *Virology* 45(2):508-515.

6. Martin MA, Bryan T, Rasheed S, & Khan AS (1981) Identification and cloning of endogenous retroviral sequences present in human DNA. *Proc Natl Acad Sci U S A* 78(8):4892-4896.

7. Tristem M (2000) Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *Journal of virology* 74(8):3715-3730.

8. Polavarapu N, Bowen NJ, & McDonald JF (2006) Identification, characterization and comparative genomics of chimpanzee endogenous retroviruses. *Genome Biol* 7(6):R51.

9. Jern P, Sperber GO, & Blomberg J (2006) Divergent patterns of recent retroviral integrations in the human and chimpanzee genomes: probable transmissions between other primates and chimpanzees. *Journal of virology* 80(3):1367-1375.

10. Villesen P, Aagaard L, Wiuf C, & Pedersen FS (2004) Identification of endogenous retroviral reading frames in the human genome. *Retrovirology* 1:32.

11. Anderssen S, Sjottem E, Svineng G, & Johansen T (1997) Comparative analyses of LTRs of the ERV-H family of primate-specific retrovirus-like elements isolated from marmoset, African green monkey, and man. *Virology* 234(1):14-30.

12. Mager DL & Freeman JD (1995) HERV-H endogenous retroviruses: presence in the New World branch but amplification in the Old World primate lineage. *Virology* 213(2):395-404.

13. Lee YN & Bieniasz PD (2007) Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog* 3(1):e10.

14. 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(4648):497-500.

15. Gao F*, et al.* (1999) Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* 397(6718):436-441.

16. Weiss RA (2006) The discovery of endogenous retroviruses. *Retrovirology* 3:67.

17. Gifford RJ*, et al.* (2008) A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. *Proc Natl Acad Sci U S A* 105(51):20362-20367.

18. Katzourakis A, Tristem M, Pybus OG, & Gifford RJ (2007) Discovery and analysis of the first endogenous lentivirus. *Proc Natl Acad Sci U S A* 104(15):6261-6265.

19. Keckesova Z, Ylinen LM, Towers GJ, Gifford RJ, & Katzourakis A (2009) Identification of a RELIK orthologue in the European hare (Lepus europaeus) reveals a minimum age of 12 million years for the lagomorph lentiviruses. *Virology* 384(1):7-11.

20. Yohn CT*, et al.* (2005) Lineage-specific expansions of retroviral insertions within the genomes of African great apes but not humans and orangutans. *PLoS Biol* 3(4):e110.

21. McCune JM*, et al.* (1988) Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. *Cell* 53(1):55-67.

22. Hallenberger S*, et al.* (1992) Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature* 360(6402):358-361.

23. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, & Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455(7209):109-113.

24. Ferrantelli F & Ruprecht RM (2002) Neutralizing antibodies against HIV -- back in the major leagues? *Curr Opin Immunol* 14(4):495-502.

25. Burton DR*, et al.* (2004) HIV vaccine design and the neutralizing antibody problem. *Nat Immunol* 5(3):233-236.

26. Buchacher A*, et al.* (1994) Generation of human monoclonal antibodies against HIV-1 proteins; electrofusion and Epstein-Barr virus transformation for peripheral blood lymphocyte immortalization. *AIDS Res Hum Retroviruses* 10(4):359-369.

27. Trkola A*, et al.* (1996) Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *Journal of virology* 70(2):1100-1108.

28. Reitter JN, Means RE, & Desrosiers RC (1998) A role for carbohydrates in immune evasion in AIDS. *Nat Med* 4(6):679-684.

29. Malenbaum SE*, et al.* (2000) The N-terminal V3 loop glycan modulates the interaction of clade A and B human immunodeficiency virus type 1 envelopes with CD4 and chemokine receptors. *Journal of virology* 74(23):11008-11016.

30. Back NK*, et al.* (1994) An N-glycan within the human immunodeficiency virus type 1 gp120 V3 loop affects virus neutralization. *Virology* 199(2):431-438.

31. Cheng-Mayer C, Brown A, Harouse J, Luciw PA, & Mayer AJ (1999) Selection for neutralization resistance of the simian/human immunodeficiency virus SHIVSF33A variant in vivo by virtue of sequence changes in the extracellular envelope glycoprotein that modify N-linked glycosylation. *Journal of virology* 73(7):5294-5300.

32. Mahalanabis M*, et al.* (2009) Continuous viral escape and selection by autologous neutralizing antibodies in drug-naive human immunodeficiency virus controllers. *Journal of virology* 83(2):662-672.

33. Richman DD, Wrin T, Little SJ, & Petropoulos CJ (2003) Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc Natl Acad Sci U S A* 100(7):4144-4149.

34. Wei X*, et al.* (2003) Antibody neutralization and escape by HIV-1. *Nature* 422(6929):307-312.

35. Goff SP (2007) Host factors exploited by retroviruses. *Nat Rev Microbiol* 5(4):253-263.

36. Cao J*, et al.* (1997) Replication and neutralization of human immunodeficiency virus type 1 lacking the V1 and V2 variable loops of the gp120 envelope glycoprotein. *Journal of virology* 71(12):9808-9812.

37. Wyatt R*, et al.* (1998) The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* 393(6686):705-711.

38. Kwong PD*, et al.* (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393(6686):648-659.

39. Rizzuto CD*, et al.* (1998) A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science* 280(5371):1949-1953.

40. Maddon PJ*, et al.* (1986) The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47(3):333-348.

41. Berson JF*, et al.* (1996) A seven-transmembrane domain receptor involved in fusion and entry of T-cell-tropic human immunodeficiency virus type 1 strains. *Journal of virology* 70(9):6288-6295.

42. Feng Y, Broder CC, Kennedy PE, & Berger EA (1996) HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272(5263):872-877.

43. Choe H*, et al.* (1996) The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 85(7):1135-1148.

44. Deng H*, et al.* (1996) Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381(6584):661-666.

45. Bjorndal A*, et al.* (1997) Coreceptor usage of primary human immunodeficiency virus type 1 isolates varies according to biological phenotype. *Journal of virology* 71(10):7478-7487.

46. Kozak SL*, et al.* (1997) CD4, CXCR-4, and CCR-5 dependencies for infections by primary patient and laboratory-adapted isolates of human immunodeficiency virus type 1. *Journal of virology* 71(2):873-882.

47. Simmons G*, et al.* (1996) Primary, syncytium-inducing human immunodeficiency virus type 1 isolates are dual-tropic and most can use either Lestr or CCR5 as coreceptors for virus entry. *Journal of virology* 70(12):8355-8360.

48. Koot M*, et al.* (1993) Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS. *Ann Intern Med* 118(9):681-688.

49. Connor RI, Sheridan KE, Ceradini D, Choe S, & Landau NR (1997) Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. *J Exp Med* 185(4):621-628.

50. Hwang SS, Boyle TJ, Lyerly HK, & Cullen BR (1991) Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science* 253(5015):71-74.

51. Shioda T, Levy JA, & Cheng-Mayer C (1991) Macrophage and T cell-line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene. *Nature* 349(6305):167-169.

52. De Jong JJ, De Ronde A, Keulen W, Tersmette M, & Goudsmit J (1992) Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support

the syncytium-inducing phenotype: analysis by single amino acid substitution. *Journal of virology* 66(11):6777-6780.

53. de Jong JJ*, et al.* (1992) Human immunodeficiency virus type 1 clones chimeric for the envelope V3 domain differ in syncytium formation and replication capacity. *Journal of virology* 66(2):757-765.

54. Briggs DR, Tuttle DL, Sleasman JW, & Goodenow MM (2000) Envelope V3 amino acid sequence predicts HIV-1 phenotype (co-receptor usage and tropism for macrophages). *AIDS* 14(18):2937-2939.

55. Fouchier RA*, et al.* (1992) Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J Virol* 66(5):3183-3187.

56. Nishimura Y*, et al.* (2005) Resting naive CD4+ T cells are massively infected and eliminated by X4-tropic simian-human immunodeficiency viruses in macaques. *Proc Natl Acad Sci U S A* 102(22):8000-8005.

57. Nishimura Y*, et al.* (2004) Highly pathogenic SHIVs and SIVs target different CD4+ T cell subsets in rhesus monkeys, explaining their divergent clinical courses. *Proc Natl Acad Sci U S A* 101(33):12324-12329.

58. Ho SH, Shek L, Gettie A, Blanchard J, & Cheng-Mayer C (2005) V3 loop-determined coreceptor preference dictates the dynamics of CD4+-T-cell loss in simian-human immunodeficiency virus-infected macaques. *J Virol* 79(19):12296-12303.

59. Ren W*, et al.* (Different tempo and anatomic location of dual-tropic and X4 virus emergence in a model of R5 simian-human immunodeficiency virus infection. *Journal of virology* 84(1):340-351.

60. Tasca S, Ho SH, & Cheng-Mayer C (2008) R5X4 viruses are evolutionary, functional, and antigenic intermediates in the pathway of a simian-human immunodeficiency virus coreceptor switch. *J Virol* 82(14):7089-7099.

61. Ho SH*, et al.* (2007) Coreceptor switch in R5-tropic simian/human immunodeficiency virus-infected macaques. *J Virol* 81(16):8621-8633.

62. Ho SH, Trunova N, Gettie A, Blanchard J, & Cheng-Mayer C (2008) Different mutational pathways to CXCR4 coreceptor switch of CCR5-using simian-human immunodeficiency virus. *Journal of virology* 82(11):5653-5656.

63. Rein A & Schultz A (1984) Different recombinant murine leukemia viruses use different cell surface receptors. *Virology* 136(1):144-152.

64. Battini JL, Danos O, & Heard JM (1998) Definition of a 14-amino-acid peptide essential for the interaction between the murine leukemia virus amphotropic envelope glycoprotein and its receptor. *Journal of virology* 72(1):428-435.

65. Battini JL, Heard JM, & Danos O (1992) Receptor choice determinants in the envelope glycoproteins of amphotropic, xenotropic, and polytropic murine leukemia viruses. *Journal of virology* 66(3):1468-1475.

66. Fass D*, et al.* (1997) Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution. *Science* 277(5332):1662-1666.

67. Albritton LM, Tseng L, Scadden D, & Cunningham JM (1989) A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection. *Cell* 57(4):659-666.

68. Kavanaugh MP*, et al.* (1994) Cell-surface receptors for gibbon ape leukemia virus and amphotropic murine retrovirus are inducible sodium-dependent phosphate symporters. *Proc Natl Acad Sci U S A* 91(15):7071-7075.

69. Miller DG, Edwards RH, & Miller AD (1994) Cloning of the cellular receptor for amphotropic murine retroviruses reveals homology to that for gibbon ape leukemia virus. *Proc Natl Acad Sci U S A* 91(1):78-82.

70. van Zeijl M*, et al.* (1994) A human amphotropic retrovirus receptor is a second member of the gibbon ape leukemia virus receptor family. *Proc Natl Acad Sci U S A* 91(3):1168-1172.

71. Miller DG & Miller AD (1994) A family of retroviruses that utilize related phosphate transporters for cell entry. *J Virol* 68(12):8270-8276.

72. Stoye JP & Coffin JM (1987) The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination. *Journal of virology* 61(9):2659-2669.

73. Battini JL, Rasko JE, & Miller AD (1999) A human cell-surface receptor for xenotropic and polytropic murine leukemia viruses: possible role in G protein-coupled signal transduction. *Proc Natl Acad Sci U S A* 96(4):1385-1390.

74. Yang YL*, et al.* (1999) Receptors for polytropic and xenotropic mouse leukaemia viruses encoded by a single gene at Rmc1. *Nat Genet* 21(2):216-219.

75. Tailor CS, Nouri A, Lee CG, Kozak C, & Kabat D (1999) Cloning and characterization of a cell surface receptor for xenotropic and polytropic murine leukemia viruses. *Proc Natl Acad Sci U S A* 96(3):927-932.

76. Elder JH*, et al.* (1977) Biochemical evidence that MCF murine leukemia viruses are envelope (env) gene recombinants. *Proc Natl Acad Sci U S A* 74(10):4676-4680.

77. Fischinger PJ, Nomura S, & Bolognesi DP (1975) A novel murine oncornavirus with dual eco- and xenotropic properties. *Proc Natl Acad Sci U S A* 72(12):5150-5155.

78. Van Hoeven NS & Miller AD (2005) Use of different but overlapping determinants in a retrovirus receptor accounts for non-reciprocal interference between xenotropic and polytropic murine leukemia viruses. *Retrovirology* 2:76.

79. Kozak CA (1985) Susceptibility of wild mouse cells to exogenous infection with xenotropic leukemia viruses: control by a single dominant locus on chromosome 1. *Journal of virology* 55(3):690-695.

80. Yan Y, Knoper RC, & Kozak CA (2007) Wild mouse variants of envelope genes of xenotropic/polytropic mouse gammaretroviruses and their XPR1 receptors elucidate receptor determinants of virus entry. *Journal of virology* 81(19):10550-10557.

81. Oliveira NM, Farrell KB, & Eiden MV (2006) In vitro characterization of a koala retrovirus. *J Virol* 80(6):3104-3107.

82. Tarlinton RE, Meers J, & Young PR (2006) Retroviral invasion of the koala genome. *Nature* 442(7098):79-81.

83. Mendoza R, Anderson MM, & Overbaugh J (2006) A putative thiamine transport protein is a receptor for feline leukemia virus subgroup A. *J Virol* 80(7):3378-3385.

84. Takeuchi Y*, et al.* (1992) Feline leukemia virus subgroup B uses the same cell surface receptor as gibbon ape leukemia virus. *J Virol* 66(2):1219-1222.

85. Anderson MM, Lauring AS, Robertson S, Dirks C, & Overbaugh J (2001) Feline Pit2 functions as a receptor for subgroup B feline leukemia viruses. *J Virol* 75(22):10563-10572.

86. Stewart MA*, et al.* (1986) Nucleotide sequences of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses. *J Virol* 58(3):825-834.

87. Quigley JG*, et al.* (2004) Identification of a human heme exporter that is essential for erythropoiesis. *Cell* 118(6):757-766.

88. Quigley JG*, et al.* (2000) Cloning of the cellular receptor for feline leukemia virus subgroup C (FeLV-C), a retrovirus that induces red cell aplasia. *Blood* 95(3):1093-1099.

89. Tailor CS, Willett BJ, & Kabat D (1999) A putative cell surface receptor for anemia-inducing feline leukemia virus subgroup C is a member of a transporter superfamily. *J Virol* 73(8):6500-6505.

90. Shalev Z*, et al.* (2009) Identification of a feline leukemia virus variant that can use THTR1, FLVCR1, and FLVCR2 for infection. *Journal of virology* 83(13):6706-6716.

91. Le Tissier P, Stoye JP, Takeuchi Y, Patience C, & Weiss RA (1997) Two sets of human-tropic pig retrovirus. *Nature* 389(6652):681-682.

92. Ericsson TA*, et al.* (2003) Identification of receptors for pig endogenous retrovirus. *Proc Natl Acad Sci U S A* 100(11):6759-6764.

93. Tailor CS, Nouri A, Zhao Y, Takeuchi Y, & Kabat D (1999) A sodium-dependent neutral-amino-acid transporter mediates infections of feline and baboon endogenous retroviruses and simian type D retroviruses. *J Virol* 73(5):4470-4474.

94. Rasko JE, Battini JL, Gottschalk RJ, Mazo I, & Miller AD (1999) The RD114/simian type D retrovirus receptor is a neutral amino acid transporter. *Proc Natl Acad Sci U S A* 96(5):2129-2134.

95. Blond JL*, et al.* (2000) An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. *Journal of virology* 74(7):3321-3329.

96. Marin M, Tailor CS, Nouri A, & Kabat D (2000) Sodium-dependent neutral amino acid transporter type 1 is an auxiliary receptor for baboon endogenous retrovirus. *Journal of virology* 74(17):8085-8093.

97. Marin M, Lavillette D, Kelly SM, & Kabat D (2003) N-linked glycosylation and sequence changes in a critical negative control region of the ASCT1 and ASCT2 neutral amino acid transporters determine their retroviral receptor functions. *Journal of virology* 77(5):2936-2945.

98. McClure MA, Johnson MS, Feng DF, & Doolittle RF (1988) Sequence comparisons of retroviral proteins: relative rates of change and general phylogeny. *Proc Natl Acad Sci U S A* 85(8):2469-2473.

99. Benit L, Dessen P, & Heidmann T (2001) Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. *J Virol* 75(23):11709-11719.

100. Blaise S, de Parseval N, Benit L, & Heidmann T (2003) Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci U S A* 100(22):13013-13018.

101. Frendo JL*, et al.* (2003) Direct involvement of HERV-W Env glycoprotein in human trophoblast cell fusion and differentiation. *Mol Cell Biol* 23(10):3566-3574.

102. Mallet F*, et al.* (2004) The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology. *Proc Natl Acad Sci U S A* 101(6):1731-1736.

103. Benit L*, et al.* (1997) Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a gag coding sequence closely related to the Fv1 restriction gene. *Journal of virology* 71(7):5652-5657.

104. Best S, Le Tissier P, Towers G, & Stoye JP (1996) Positional cloning of the mouse retrovirus restriction gene Fv1. *Nature* 382(6594):826-829.

105. Lilly F (1967) Susceptibility to two strains of Friend leukemia virus in mice. *Science* 155(761):461-462.

106. Kozak CA & Chakraborti A (1996) Single amino acid changes in the murine leukemia virus capsid protein gene define the target of Fv1 resistance. *Virology* 225(2):300-305.

107. Kozak CA (1985) Analysis of wild-derived mice for Fv-1 and Fv-2 murine leukemia virus restriction loci: a novel wild mouse Fv-1 allele responsible for lack of host range restriction. *Journal of virology* 55(2):281-285.

108. Hatziioannou T, Perez-Caballero D, Yang A, Cowan S, & Bieniasz PD (2004) Retrovirus resistance factors Ref1 and Lv1 are species-specific variants of TRIM5alpha. *Proc Natl Acad Sci U S A* 101(29):10774-10779.

109. Perron MJ*, et al.* (2004) TRIM5alpha mediates the postentry block to N-tropic murine leukemia viruses in human cells. *Proc Natl Acad Sci U S A* 101(32):11827-11832.

110. Yap MW, Nisole S, Lynch C, & Stoye JP (2004) Trim5alpha protein restricts both HIV-1 and murine leukemia virus. *Proc Natl Acad Sci U S A* 101(29):10786-10791.

111. Keckesova Z, Ylinen LM, & Towers GJ (2004) The human and African green monkey TRIM5alpha genes encode Ref1 and Lv1 retroviral restriction factor activities. *Proc Natl Acad Sci U S A* 101(29):10780-10785.

112. Stremlau M*, et al.* (2004) The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature* 427(6977):848-853.

113. Mortuza GB*, et al.* (2004) High-resolution structure of a retroviral capsid hexameric amino-terminal domain. *Nature* 431(7007):481-485.

114. Lassaux A, Sitbon M, & Battini JL (2005) Residues in the murine leukemia virus capsid that differentially govern resistance to mouse Fv1 and human Ref1 restrictions. *Journal of virology* 79(10):6560-6564.

115. Kratovac Z*, et al.* (2008) Primate lentivirus capsid sensitivity to TRIM5 proteins. *J Virol* 82(13):6772-6777.

116. Reymond A*, et al.* (2001) The tripartite motif family identifies cell compartments. *Embo J* 20(9):2140-2151.

117. Nakayama EE, Miyoshi H, Nagai Y, & Shioda T (2005) A specific region of 37 amino acid residues in the SPRY (B30.2) domain of African green monkey TRIM5alpha determines species-specific restriction of simian immunodeficiency virus SIVmac infection. *Journal of virology* 79(14):8870-8877.

118. Sebastian S & Luban J (2005) TRIM5alpha selectively binds a restriction-sensitive retroviral capsid. *Retrovirology* 2:40.

119. Stremlau M, Perron M, Welikala S, & Sodroski J (2005) Species-specific variation in the B30.2(SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. *Journal of virology* 79(5):3139-3145.

120. Yap MW, Nisole S, & Stoye JP (2005) A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. *Curr Biol* 15(1):73-78.

121. Perez-Caballero D, Hatziioannou T, Yang A, Cowan S, & Bieniasz PD (2005) Human tripartite motif 5alpha domains responsible for retrovirus restriction activity and specificity. *Journal of virology* 79(14):8969-8978.

122. Diaz-Griffero F*, et al.* (2006) Rapid turnover and polyubiquitylation of the retroviral restriction factor TRIM5. *Virology* 349(2):300-315.

123. Yamauchi K, Wada K, Tanji K, Tanaka M, & Kamitani T (2008) Ubiquitination of E3 ubiquitin ligase TRIM5 alpha and its potential role. *FEBS J* 275(7):1540-1555.

124. Rold CJ & Aiken C (2008) Proteasomal degradation of TRIM5alpha during retrovirus restriction. *PLoS Pathog* 4(5):e1000074.

125. Stremlau M*, et al.* (2006) Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc Natl Acad Sci U S A* 103(14):5514-5519.

126. Anderson JL*, et al.* (2006) Proteasome inhibition reveals that a functional preintegration complex intermediate can be generated during restriction by diverse TRIM5 proteins. *Journal of virology* 80(19):9754-9760.

127. Wu X, Anderson JL, Campbell EM, Joseph AM, & Hope TJ (2006) Proteasome inhibitors uncouple rhesus TRIM5alpha restriction of HIV-1 reverse transcription and infection. *Proc Natl Acad Sci U S A* 103(19):7465-7470.

128. Perez-Caballero D, Hatziioannou T, Zhang F, Cowan S, & Bieniasz PD (2005) Restriction of human immunodeficiency virus type 1 by TRIM-CypA occurs with rapid kinetics and independently of cytoplasmic bodies, ubiquitin, and proteasome activity. *Journal of virology* 79(24):15567-15572.

129. Franke EK, Yuan HE, & Luban J (1994) Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* 372(6504):359-362.

130. Luban J, Bossolt KL, Franke EK, Kalpana GV, & Goff SP (1993) Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. *Cell* 73(6):1067-1078.

131. Thali M*, et al.* (1994) Functional association of cyclophilin A with HIV-1 virions. *Nature* 372(6504):363-365.

132. Hatziioannou T, Perez-Caballero D, Cowan S, & Bieniasz PD (2005) Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells. *J Virol* 79(1):176-183.

133. Sokolskaja E, Sayah DM, & Luban J (2004) Target cell cyclophilin A modulates human immunodeficiency virus type 1 infectivity. *J Virol* 78(23):12800-12808.

134. Aberham C, Weber S, & Phares W (1996) Spontaneous mutations in the human immunodeficiency virus type 1 gag gene that affect viral replication in the presence of cyclosporins. *Journal of virology* 70(6):3536-3544.

135. Braaten D*, et al.* (1996) Cyclosporine A-resistant human immunodeficiency virus type 1 mutants demonstrate that Gag encodes the functional target of cyclophilin A. *Journal of virology* 70(8):5170-5176.

136. Berthoux L, Sebastian S, Sokolskaja E, & Luban J (2005) Cyclophilin A is required for TRIM5{alpha}-mediated resistance to HIV-1 in Old World monkey cells. *Proc Natl Acad Sci U S A* 102(41):14849-14853.

137. Kootstra NA, Munk C, Tonnu N, Landau NR, & Verma IM (2003) Abrogation of postentry restriction of HIV-1-based lentiviral vector transduction in simian cells. *Proc Natl Acad Sci U S A* 100(3):1298-1303.

138. Towers GJ*, et al.* (2003) Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. *Nat Med* 9(9):1138-1143.

139. Wilson SJ*, et al.* (2008) Independent evolution of an antiviral TRIMCyp in rhesus macaques. *Proceedings of the National Academy of Sciences of the United States of America* 105(9):3557-3562.

140. Virgen CA, Kratovac Z, Bieniasz PD, & Hatziioannou T (2008) Independent genesis of chimeric TRIM5-cyclophilin proteins in two primate species.

Proceedings of the National Academy of Sciences of the United States of America 105(9):3563-3568.

141. Brennan G, Kozyrev Y, & Hu SL (2008) TRIMCyp expression in Old World primates Macaca nemestrina and Macaca fascicularis. *Proceedings of the National Academy of Sciences of the United States of America* 105(9):3569-3574.

142. Liao CH, Kuang YQ, Liu HL, Zheng YT, & Su B (2007) A novel fusion gene, TRIM5-Cyclophilin A in the pig-tailed macaque determines its susceptibility to HIV-1 infection. *AIDS (London, England)* 21 Suppl 8:S19-26.

143. Sayah DM, Sokolskaja E, Berthoux L, & Luban J (2004) Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* 430(6999):569-573.

144. Nisole S, Lynch C, Stoye JP, & Yap MW (2004) A Trim5-cyclophilin A fusion protein found in owl monkey kidney cells can restrict HIV-1. *Proceedings of the National Academy of Sciences of the United States of America* 101(36):13324-13328.

145. 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(8):2832-2837.

146. Song B*, et al.* (2005) The B30.2(SPRY) domain of the retroviral restriction factor TRIM5alpha exhibits lineage-specific length and sequence variation in primates. *Journal of virology* 79(10):6111-6121.

147. Goff SP (2007) Retroviridae: The Retroviruses and Their Replication. *Fields Virology*, ed PM KDH (Lippincott Williams & Wilkins, Philidelphia), Fifth Ed, pp 1999-2070.

148. Coffin JM (1979) Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *J Gen Virol* 42(1):1-26.

149. Gilboa E, Mitra SW, Goff S, & Baltimore D (1979) A detailed model of reverse transcription and tests of crucial aspects. *Cell* 18(1):93-100.

150. Haseltine WA, Coffin JM, & Hageman TC (1979) Structure of products of the Moloney murine leukemia virus endogenous DNA polymerase reaction. *Journal of virology* 30(1):375-383.

151. Zheng YH*, et al.* (2004) Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *Journal of virology* 78(11):6073-6076.

152. 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(6898):646-650.

153. Liddament MT, Brown WL, Schumacher AJ, & Harris RS (2004) APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. *Curr Biol* 14(15):1385-1391.

154. Wiegand HL, Doehle BP, Bogerd HP, & Cullen BR (2004) A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *Embo J* 23(12):2451-2458.

155. Zennou V & Bieniasz PD (2006) Comparative analysis of the antiretroviral activity of APOBEC3G and APOBEC3F from primates. *Virology* 349(1):31-40.

156. Virgen CA & Hatziioannou T (2007) Antiretroviral activity and Vif sensitivity of rhesus macaque APOBEC3 proteins. *Journal of virology* 81(24):13932-13937.

157. 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(3):249-259.

158. OhAinle M, Kerns JA, Malik HS, & Emerman M (2006) Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H. *Journal of virology* 80(8):3853-3862.

159. Dang Y, Wang X, Esselman WJ, & Zheng YH (2006) Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *Journal of virology* 80(21):10522-10533.

160. Bishop KN*, et al.* (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr Biol* 14(15):1392-1396.

161. Yu Q*, et al.* (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J Biol Chem* 279(51):53379-53386.

162. Navarro F*, et al.* (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology* 333(2):374-386.

163. 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(2):163-168.

164. Burnett A & Spearman P (2007) APOBEC3G multimers are recruited to the plasma membrane for packaging into human immunodeficiency virus type 1 virus-like particles in an RNA-dependent process requiring the NC basic linker. *J Virol* 81(10):5000-5013.

165. Zennou V, Perez-Caballero D, Gottlinger H, & Bieniasz PD (2004) APOBEC3G incorporation into human immunodeficiency virus type 1 particles. *Journal of virology* 78(21):12058-12061.

166. Khan MA*, et al.* (2007) Analysis of the contribution of cellular and viral RNA to the packaging of APOBEC3G into HIV-1 virions. *Retrovirology* 4:48.

167. Khan MA*, et al.* (2005) Viral RNA is required for the association of APOBEC3G with human immunodeficiency virus type 1 nucleoprotein complexes. *Journal of virology* 79(9):5870-5874.

168. Bogerd HP & Cullen BR (2008) Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation. *RNA* 14(6):1228-1236.

169. Mangeat B*, et al.* (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424(6944):99-103.

170. Harris RS*, et al.* (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell* 113(6):803-809.

171. Yu Q*, et al.* (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nature structural & molecular biology* 11(5):435-442.

172. Suspene R, Rusniok C, Vartanian JP, & Wain-Hobson S (2006) Twin gradients in APOBEC3 edited HIV-1 DNA reflect the dynamics of lentiviral replication. *Nucleic acids research* 34(17):4677-4684.

173. Shirakawa K*, et al.* (2006) Ubiquitination of APOBEC3 proteins by the Vif-Cullin5-ElonginB-ElonginC complex. *Virology* 344(2):263-266.

174. Yu X*, et al.* (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302(5647):1056-1060.

175. Hatziioannou T*, et al.* (2006) Generation of simian-tropic HIV-1 by restriction factor evasion. *Science* 314(5796):95.

176. Sawyer SL, Emerman M, & Malik HS (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol* 2(9):E275.

177. Harris RS & Liddament MT (2004) Retroviral restriction by APOBEC proteins. *Nat Rev Immunol* 4(11):868-877.

178. Courcoul M*, et al.* (1995) Peripheral blood mononuclear cells produce normal amounts of defective Vif- human immunodeficiency virus type 1 particles which are restricted for the preretrotranscription steps. *Journal of virology* 69(4):2068-2074.

179. von Schwedler U, Song J, Aiken C, & Trono D (1993) Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. *Journal of virology* 67(8):4945-4955.

180. Kaiser SM & Emerman M (2006) Uracil DNA glycosylase is dispensable for human immunodeficiency virus type 1 replication and does not contribute to the antiviral effects of the cytidine deaminase Apobec3G. *Journal of virology* 80(2):875-882.

181. Holmes RK, Koning FA, Bishop KN, & Malim MH (2007) APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J Biol Chem* 282(4):2587-2595.

182. Bishop KN, Holmes RK, & Malim MH (2006) Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. *Journal of virology* 80(17):8450-8458.

183. Newman EN*, et al.* (2005) Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Curr Biol* 15(2):166-170.

184. Opi S*, et al.* (2006) Monomeric APOBEC3G is catalytically active and has antiviral activity. *Journal of virology* 80(10):4673-4682.

185. Schumacher AJ, Hache G, Macduff DA, Brown WL, & Harris RS (2008) The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. *J Virol* 82(6):2652-2660.

186. Miyagi E*, et al.* (2007) Enzymatically active APOBEC3G is required for efficient inhibition of human immunodeficiency virus type 1. *Journal of virology* 81(24):13346-13353.

187. Browne EP, Allers C, & Landau NR (2009) Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent. *Virology* 387(2):313-321.

188. Mbisa JL*, et al.* (2007) Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J Virol* 81(13):7099-7110.

189. Guo F, Cen S, Niu M, Saadatmand J, & Kleiman L (2006) Inhibition of formula-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. *Journal of virology* 80(23):11710-11722.

190. Guo F*, et al.* (2007) The Interaction of APOBEC3G with Human Immunodeficiency Virus Type 1 Nucleocapsid Inhibits tRNA3Lys Annealing to Viral RNA. *Journal of virology* 81(20):11322-11331.

191. Li XY, Guo F, Zhang L, Kleiman L, & Cen S (2007) APOBEC3G inhibits DNA strand transfer during HIV-1 reverse transcription. *J Biol Chem*.

192. Janini M, Rogers M, Birx DR, & McCutchan FE (2001) Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. *Journal of virology* 75(17):7973-7986.

193. Kieffer TL*, et al.* (2005) G-->A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *Journal of virology* 79(3):1975-1980.

194. Bishop KN, Holmes RK, Sheehy AM, & Malim MH (2004) APOBEC-mediated editing of viral RNA. *Science* 305(5684):645.

195. Tian C*, et al.* (2006) Differential requirement for conserved tryptophans in human immunodeficiency virus type 1 Vif for the selective suppression of APOBEC3G and APOBEC3F. *Journal of virology* 80(6):3112-3115.

196. Russell RA & Pathak VK (2007) Identification of two distinct human immunodeficiency virus type 1 Vif determinants critical for interactions with human APOBEC3G and APOBEC3F. *Journal of virology* 81(15):8201-8210.

197. Simon V*, et al.* (2005) Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS Pathog* 1(1):e6.

198. Eigen M (1993) The origin of genetic information: viruses as models. *Gene* 135(1-2):37-47.

199. Hu WS & Temin HM (1990) Retroviral recombination and reverse transcription. *Science* 250(4985):1227-1233.

200. Hu WS & Temin HM (1990) Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc Natl Acad Sci U S A* 87(4):1556-1560.

201. Stuhlmann H & Berg P (1992) Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *Journal of virology* 66(4):2378-2388.

202. Zhang J & Temin HM (1993) Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. *Science* 259(5092):234-238.

203. Rhodes T, Wargo H, & Hu WS (2003) High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication. *Journal of virology* 77(20):11193-11200.

204. Fisher RA (1930) *The genetical theory of natural selection* (Oxford University Press, Oxford).

205. Muller HJ (1964) The Relation of Recombination to Mutational Advance. *Mutat Res* 106:2-9.

206. Lehman N (2003) A case for the extreme antiquity of recombination. *J Mol Evol* 56(6):770-777.

207. Soll SJ, Diaz Arenas C, & Lehman N (2007) Accumulation of deleterious mutations in small abiotic populations of RNA. *Genetics* 175(1):267-275.

208. Salazar-Gonzalez JF*, et al.* (2008) Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *Journal of virology* 82(8):3952-3970.

209. Salazar-Gonzalez JF*, et al.* (2009) Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J Exp Med* 206(6):1273-1289.

210. Wood N*, et al.* (2009) HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. *PLoS Pathog* 5(5):e1000414.

211. Mulder LC, Harari A, & Simon V (2008) Cytidine deamination induced HIV-1 drug resistance. *Proc Natl Acad Sci U S A* 105(14):5501-5506.

212. Bowerman B, Brown PO, Bishop JM, & Varmus HE (1989) A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev* 3(4):469-478.

213. Miller MD, Farnet CM, & Bushman FD (1997) Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *Journal of virology* 71(7):5382-5390.

214. Goff SP (1992) Genetics of retroviral integration. *Annu Rev Genet* 26:527-544.

215. Shoemaker C*, et al.* (1980) Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: implications for retrovirus integration. *Proc Natl Acad Sci U S A* 77(7):3932-3936.

216. Speck NA & Baltimore D (1987) Six distinct nuclear factors interact with the 75-base-pair repeat of the Moloney murine leukemia virus enhancer. *Mol Cell Biol* 7(3):1101-1110.

217. Ryden TA & Beemon K (1989) Avian retroviral long terminal repeats bind CCAAT/enhancer-binding protein. *Mol Cell Biol* 9(3):1155-1164.

218. Jakobovits A, Smith DH, Jakobovits EB, & Capon DJ (1988) A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. *Mol Cell Biol* 8(6):2555-2561.

219. Rice AP & Mathews MB (1988) Transcriptional but not translational regulation of HIV-1 by the tat gene product. *Nature* 332(6164):551-553.

220. Laspia MF, Rice AP, & Mathews MB (1989) HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. *Cell* 59(2):283-292.

221. Muesing MA, Smith DH, & Capon DJ (1987) Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. *Cell* 48(4):691-701.

222. Hauber J, Perkins A, Heimer EP, & Cullen BR (1987) Trans-activation of human immunodeficiency virus gene expression is mediated by nuclear events. *Proc Natl Acad Sci U S A* 84(18):6364-6368.

223. Fisher AG*, et al.* (1986) The trans-activator gene of HTLV-III is essential for virus replication. *Nature* 320(6060):367-371.

224. Dayton AI, Sodroski JG, Rosen CA, Goh WC, & Haseltine WA (1986) The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* 44(6):941-947.

225. Fujinaga K*, et al.* (1998) The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat. *Journal of virology* 72(9):7154-7159.

226. Isel C & Karn J (1999) Direct evidence that HIV-1 Tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation. *J Mol Biol* 290(5):929-941.

227. Zhou M*, et al.* (2000) Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol Cell Biol* 20(14):5077-5086.

228. Parada CA & Roeder RG (1996) Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature* 384(6607):375-378.

229. Purcell DF & Martin MA (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *Journal of virology* 67(11):6365-6378.

230. Feinberg MB, Jarrett RF, Aldovini A, Gallo RC, & Wong-Staal F (1986) HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell* 46(6):807-817.

231. Knight DM, Flomerfelt FA, & Ghrayeb J (1987) Expression of the art/trs protein of HIV and study of its role in viral envelope synthesis. *Science* 236(4803):837-840.

232. Hadzopoulou-Cladaras M*, et al.* (1989) The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region. *Journal of virology* 63(3):1265-1274.

233. Rosen CA, Terwilliger E, Dayton A, Sodroski JG, & Haseltine WA (1988) Intragenic cis-acting art gene-responsive sequences of the human immunodeficiency virus. *Proc Natl Acad Sci U S A* 85(7):2071-2075.

234. Felber BK, Hadzopoulou-Cladaras M, Cladaras C, Copeland T, & Pavlakis GN (1989) rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc Natl Acad Sci U S A* 86(5):1495-1499.

235. Hammarskjold ML*, et al.* (1989) Regulation of human immunodeficiency virus env expression by the rev gene product. *Journal of virology* 63(5):1959-1966.

236. Sodroski J*, et al.* (1986) A second post-transcriptional trans-activator gene required for HTLV-III replication. *Nature* 321(6068):412-417.

237. Yoshinaka Y, Katoh I, Copeland TD, & Oroszlan S (1985) Translational readthrough of an amber termination codon during synthesis of feline leukemia virus protease. *Journal of virology* 55(3):870-873.

238. Yoshinaka Y, Katoh I, Copeland TD, & Oroszlan S (1985) Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. *Proc Natl Acad Sci U S A* 82(6):1618-1622.

239. Jacks T & Varmus HE (1985) Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* 230(4731):1237-1242.

240. Jouvenet N, Bieniasz PD, & Simon SM (2008) Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* 454(7201):236-240.

241. Jouvenet N*, et al.* (2006) Plasma membrane is the site of productive HIV-1 particle assembly. *PLoS Biol* 4(12):e435.

242. Jouvenet N, Simon SM, & Bieniasz PD (2009) Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. *Proc Natl Acad Sci U S A* 106(45):19114-19119.

243. Bieniasz PD (2006) Late budding domains and host proteins in enveloped virus release. *Virology* 344(1):55-63.

244. Morita E & Sundquist WI (2004) Retrovirus budding. *Annu Rev Cell Dev Biol* 20:395-425.

245. Katoh I*, et al.* (1985) Murine leukemia virus maturation: protease region required for conversion from "immature" to "mature" core form and for virus infectivity. *Virology* 145(2):280-292.

246. Crawford S & Goff SP (1985) A deletion mutation in the 5' part of the pol gene of Moloney murine leukemia virus blocks proteolytic processing of the gag and pol polyproteins. *Journal of virology* 53(3):899-907.

247. Van Damme N*, et al.* (2008) The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* 3(4):245-252.

248. Neil SJ, Zang T, & Bieniasz PD (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* 451(7177):425-430.

249. Perez-Caballero D*, et al.* (2009) Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell* 139(3):499-511.

250. Jouvenet N*, et al.* (2009) Broad-spectrum inhibition of retroviral and filoviral particle release by tetherin. *J Virol* 83(4):1837-1844.

251. Weiss RA, Friis RR, Katz E, & Vogt PK (1971) Induction of avian tumor viruses in normal cells by physical and chemical carcinogens. *Virology* 46(3):920-938.

252. Payne LN & Chubb RC (1968) Studies on the nature and genetic control of an antigen in normal chick embryos which reacts in the COFAL test. *J Gen Virol* 3(3):379-391.

253. Lander ES*, et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860-921.

254. Dewannieux M*, et al.* (2006) Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res* 16(12):1548-1556.

255. Perez-Caballero D, Soll SJ, & Bieniasz PD (2008) Evidence for restriction of ancient primate gammaretroviruses by APOBEC3 but not TRIM5alpha proteins. *PLoS Pathog* 4(10):e1000181.

256. Soneoka Y*, et al.* (1995) A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res* 23(4):628-633.

257. Bainbridge JW*, et al.* (2001) In vivo gene transfer to the mouse eye using an HIV-based lentiviral vector; efficient long-term transduction of corneal endothelium and retinal pigment epithelium. *Gene Ther* 8(21):1665-1668.

258. Zhang YJ*, et al.* (2002) Envelope-dependent, cyclophilin-independent effects of glycosaminoglycans on human immunodeficiency virus type 1 attachment and infection. *J Virol* 76(12):6332-6343.

259. Bieniasz PD & Cullen BR (2000) Multiple blocks to human immunodeficiency virus type 1 replication in rodent cells. *J Virol* 74(21):9868-9877.

260. Kaiser SM, Malik HS, & Emerman M (2007) Restriction of an extinct retrovirus by the human TRIM5alpha antiviral protein. *Science* 316(5832):1756-1758.

261. Mariani R*, et al.* (2003) Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* 114(1):21-31.

262. Neil SJ, Eastman SW, Jouvenet N, & Bieniasz PD (2006) HIV-1 Vpu promotes release and prevents endocytosis of nascent retrovirus particles from the plasma membrane. *PLoS Pathog* 2(5):e39.

263. Baba M*, et al.* (1988) Mechanism of inhibitory effect of dextran sulfate and heparin on replication of human immunodeficiency virus in vitro. *Proc Natl Acad Sci U S A* 85(16):6132-6136.

264. Bird AP (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* 8(7):1499-1504.

265. Duncan BK & Miller JH (1980) Mutagenic deamination of cytosine residues in DNA. *Nature* 287(5782):560-561.

266. Popp C*, et al.* (Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 463(7284):1101-1105.

267. Bhutani N*, et al.* (Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 463(7284):1042-1047.

268. Jern P, Stoye JP, & Coffin JM (2007) Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses. *PLoS Genet* 3(10):2014-2022.

269. Jarmuz A*, et al.* (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* 79(3):285-296.

270. Lee YN, Malim MH, & Bieniasz PD (2008) Hypermutation of an ancient human retrovirus by APOBEC3G. *Journal of virology* 82(17):8762-8770.

271. Kijak GH*, et al.* (2008) Variable contexts and levels of hypermutation in HIV-1 proviral genomes recovered from primary peripheral blood mononuclear cells. *Virology* 376(1):101-111.

272. Derse D, Hill SA, Princler G, Lloyd P, & Heidecker G (2007) Resistance of human T cell leukemia virus type 1 to APOBEC3G restriction is mediated by elements in nucleocapsid. *Proc Natl Acad Sci U S A* 104(8):2915-2920.

273. Lellek H*, et al.* (2000) Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex. *J Biol Chem* 275(26):19848-19856.

274. Mehta A, Kinter MT, Sherman NE, & Driscoll DM (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol* 20(5):1846-1854.

275. Hache G, Shindo K, Albin JS, & Harris RS (2008) Evolution of HIV-1 isolates that use a novel Vif-independent mechanism to resist restriction by human APOBEC3G. *Curr Biol* 18(11):819-824.

276. Lawrence J*, et al.* (1999) Clinical resistance patterns and responses to two sequential protease inhibitor regimens in saquinavir and reverse transcriptase inhibitor-experienced persons. *J Infect Dis* 179(6):1356-1364.

277. Patick AK*, et al.* (1998) Genotypic and phenotypic characterization of human immunodeficiency virus type 1 variants isolated from patients treated with the protease inhibitor nelfinavir. *Antimicrob Agents Chemother* 42(10):2637-2644.

278. Kellam P, Boucher CA, & Larder BA (1992) Fifth mutation in human immunodeficiency virus type 1 reverse transcriptase contributes to the development of high-level resistance to zidovudine. *Proc Natl Acad Sci U S A* 89(5):1934-1938.

279. Kemp SD*, et al.* (1998) A novel polymorphism at codon 333 of human immunodeficiency virus type 1 reverse transcriptase can facilitate dual resistance to zidovudine and L-2',3'-dideoxy-3'-thiacytidine. *J Virol* 72(6):5093-5098.

280. Larder BA, Coates KE, & Kemp SD (1991) Zidovudine-resistant human immunodeficiency virus selected by passage in cell culture. *J Virol* 65(10):5232-5236.

281. Larder BA & Kemp SD (1989) Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* 246(4934):1155-1158.

282. Larder BA, Kemp SD, & Purifoy DJ (1989) Infectious potential of human immunodeficiency virus type 1 reverse transcriptase mutants with altered inhibitor sensitivity. *Proc Natl Acad Sci U S A* 86(13):4803-4807.

283. Miller V, *et al.* (1998) Dual resistance to zidovudine and lamivudine in patients treated with zidovudine-lamivudine combination therapy: association with therapy failure. *J Infect Dis* 177(6):1521-1532.

284. Blanco JL, *et al.* (2005) Evolution of resistance mutations pattern in HIV-1-infected patients during intensification therapy with a boosted protease inhibitor. *AIDS* 19(8):829-831.

285. Fan N, *et al.* (1996) A drug resistance mutation in the inhibitor binding pocket of human immunodeficiency virus type 1 reverse transcriptase impairs DNA synthesis and RNA degradation. *Biochemistry* 35(30):9737-9745.

286. Gilbert PB, *et al.* (2000) Comparative analysis of HIV type 1 genotypic resistance across antiretroviral trial treatment regimens. *AIDS Res Hum Retroviruses* 16(14):1325-1336.

287. Hanna GJ, *et al.* (2000) Patterns of resistance mutations selected by treatment of human immunodeficiency virus type 1 infection with zidovudine, didanosine, and nevirapine. *J Infect Dis* 181(3):904-911.

288. Huang W, Gamarnik A, Limoli K, Petropoulos CJ, & Whitcomb JM (2003) Amino acid substitutions at position 190 of human immunodeficiency virus type 1 reverse transcriptase increase susceptibility to delavirdine and impair virus replication. *J Virol* 77(2):1512-1523.

289. Matsumi S, *et al.* (2003) Pathways for the emergence of multi-dideoxynucleoside-resistant HIV-1 variants. *AIDS* 17(8):1127-1137.

290. Nissley DV, *et al.* (2005) Sensitive phenotypic detection of minor drug-resistant human immunodeficiency virus type 1 reverse transcriptase variants. *J Clin Microbiol* 43(11):5696-5704.

291. Ochoa de Echaguen A, *et al.* (2005) Genotypic and phenotypic resistance patterns at virological failure in a simplification trial with nevirapine, efavirenz or abacavir. *AIDS* 19(13):1385-1391.

292. Olmsted RA, *et al.* (1996) (Alkylamino) piperidine bis(heteroaryl)piperizine analogs are potent, broad-spectrum nonnucleoside reverse transcriptase inhibitors of drug-resistant isolates of human immunodeficiency virus type 1 (HIV-1) and select for drug-resistant variants of HIV-1IIIB with reduced replication phenotypes. *J Virol* 70(6):3698-3705.

293. Petropoulos CJ, *et al.* (2000) A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother* 44(4):920-928.

294. Richman DD, *et al.* (1994) Nevirapine resistance mutations of human immunodeficiency virus type 1 selected during therapy. *J Virol* 68(3):1660-1666.

295. Cao Y, Qin L, Zhang L, Safrit J, & Ho DD (1995) Virologic and immunologic characterization of long-term survivors of human immunodeficiency virus type 1 infection. *N Engl J Med* 332(4):201-208.

296. Zhou B & Gitschier J (1997) hCTR1: a human gene for copper uptake identified by complementation in yeast. *Proc Natl Acad Sci U S A* 94(14):7481-7486.

297. Petris MJ, Smith K, Lee J, & Thiele DJ (2003) Copper-stimulated endocytosis and degradation of the human copper transporter, hCtr1. *J Biol Chem* 278(11):9639-9646.

298. Eisses JF & Kaplan JH (2002) Molecular characterization of hCTR1, the human copper uptake protein. *J Biol Chem* 277(32):29162-29171.

299. Klomp AE*, et al.* (2003) The N-terminus of the human copper transporter 1 (hCTR1) is localized extracellularly, and interacts with itself. *Biochem J* 370(Pt 3):881-889.

300. Aller SG, Eng ET, De Feo CJ, & Unger VM (2004) Eukaryotic CTR copper uptake transporters require two faces of the third transmembrane domain for helix packing, oligomerization, and function. *J Biol Chem* 279(51):53435-53441.

301. Eisses JF & Kaplan JH (2005) The mechanism of copper uptake mediated by human CTR1: a mutational analysis. *J Biol Chem* 280(44):37159-37168.

302. De Feo CJ, Aller SG, Siluvai GS, Blackburn NJ, & Unger VM (2009) Three-dimensional structure of the human copper transporter hCTR1. *Proc Natl Acad Sci U S A* 106(11):4237-4242.

303. Puig S, Lee J, Lau M, & Thiele DJ (2002) Biochemical and genetic analyses of yeast and human high affinity copper transporters suggest a conserved mechanism for copper uptake. *J Biol Chem* 277(29):26021-26030.

304. Rulli SJ, Jr.*, et al.* (2008) Interactions of murine APOBEC3 and human APOBEC3G with murine leukemia viruses. *Journal of virology* 82(13):6566-6575.

305. Newman RM*, et al.* (2006) Balancing selection and the evolution of functional polymorphism in Old World monkey TRIM5alpha. *Proc Natl Acad Sci U S A* 103(50):19134-19139.

306. Dodding MP, Bock M, Yap MW, & Stoye JP (2005) Capsid processing requirements for abrogation of Fv1 and Ref1 restriction. *Journal of virology* 79(16):10571-10577.

307. Hatziioannou T*, et al.* (2009) A macaque model of HIV-1 infection. *Proc Natl Acad Sci U S A* 106(11):4425-4429.

308. Eisses JF, Chi Y, & Kaplan JH (2005) Stable plasma membrane levels of hCTR1 mediate cellular copper uptake. *J Biol Chem* 280(10):9635-9639.

309. Klomp AE, Tops BB, Van Denberg IE, Berger R, & Klomp LW (2002) Biochemical characterization and subcellular localization of human copper transporter 1 (hCTR1). *Biochem J* 364(Pt 2):497-505.

310. Molloy SA & Kaplan JH (2009) Copper-dependent recycling of hCTR1, the human high affinity copper transporter. *J Biol Chem* 284(43):29704-29713.

311. Lee J, Pena MM, Nose Y, & Thiele DJ (2002) Biochemical characterization of the human copper transporter Ctr1. *J Biol Chem* 277(6):4380-4387.

312. Wang H, Kavanaugh MP, & Kabat D (1994) A critical site in the cell surface receptor for ecotropic murine retroviruses required for amino acid transport but not for viral reception. *Virology* 202(2):1058-1060.

313. Su AI*, et al.* (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* 101(16):6062-6067.

314. Barbulescu M*, et al.* (2001) A HERV-K provirus in chimpanzees, bonobos and gorillas, but not humans. *Curr Biol* 11(10):779-783.

315. Turner G*, et al.* (2001) Insertional polymorphisms of full-length endogenous retroviruses in humans. *Curr Biol* 11(19):1531-1535.

316. Wolf D & Goff SP (2009) Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* 458(7242):1201-1204.