Evidence linking APOBEC3B genesis and evolution of innate immune antagonism by gamma-herpesvirus ribonucleotide reductases

Short title: APOBEC3B antagonism by gamma-herpesviruses

Sofia N. Moraes[1], Jordan T. Becker[1], Seyed Arad Moghadasi[1], Nadine M. Shaban[1], Ashley A. Auerbach[1,2], Adam Z. Cheng[1], Reuben S. Harris[1,2,3,*]

[1] Department of Biochemistry, Molecular Biology, and Biophysics, Institute for Molecular Virology, Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455, USA.

[2] Department of Biochemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, TX 78229, USA.

[3] Howard Hughes Medical Institute, University of Texas Health San Antonio, San Antonio, TX 78229, USA.

* Corresponding author. Email: rsh@uthscsa.edu

Key words: APOBEC3B; evolutionary arms race; herpesvirus; host-pathogen interaction; innate immunity; virus restriction factor; ribonucleotide reductase; virus evolution

Abstract

Viruses have evolved diverse mechanisms to antagonize host immunity such as direct inhibition and relocalization of cellular APOBEC3B (A3B) by the ribonucleotide reductase (RNR) of Epstein-Barr virus. Here, we investigate the mechanistic conservation and evolutionary origin of this innate immune counteraction strategy. First, we find that human gamma-herpesvirus RNRs engage A3B via largely distinct surfaces. Second, we show that RNR-mediated enzymatic inhibition and relocalization of A3B depend upon binding to different regions of the catalytic domain. Third, we show that the capability of viral RNRs to antagonize A3B is conserved among gamma-herpesviruses that infect humans and Old World monkeys that encode this enzyme but absent in homologous viruses that infect New World monkeys that naturally lack the *A3B* gene. Finally, we reconstruct the ancestral primate A3B protein and demonstrate that it is active and similarly engaged by the RNRs from viruses that infect humans and Old World monkeys but not by the RNRs from viruses that infect New World monkeys. These results combine to indicate that the birth of *A3B* at a critical branchpoint in primate evolution may have been a driving force in selecting for an ancestral gamma-herpesvirus with an expanded RNR functionality through counteraction of this antiviral enzyme.

Introduction

Cell-intrinsic innate immune proteins, or 'restriction factors', contribute to the first line of defenses against incoming viruses (reviewed by [1]). A prototypical example is the APOBEC3 (A3) family of single-stranded (ss)DNA cytosine deaminases, which provide broad and overlapping protection against viral infections (reviewed by [2, 3]). Classically, A3s exert antiviral activity by catalyzing cytosine deamination in exposed ssDNA replication intermediates, often resulting in lethal mutagenesis of viral genomes. Humans encode seven A3 enzymes (A3A-D, A3F-H), which have been implicated in the restriction of retroviruses (HIV and HTLV) [4-8], endogenous retroviruses (HERV) [9, 10], hepadnaviruses (HBV) [11, 12], papillomaviruses (HPV) [13], polyomaviruses (BK) [14], and most recently the gamma-herpesvirus Epstein-Barr virus (EBV [15]). It is difficult to predict which subset of A3 enzymes has the potential to restrict a given virus, but differences in the subcellular localization properties of these enzymes provide useful insights into their antiviral targets. For instance, cytoplasmic A3s like A3G and A3F potently restrict HIV-1 by binding to viral genomic RNAs in the cytoplasm, packaging into budding virions at the cell membrane, and deaminating viral ssDNA intermediates during reverse transcription [16-18]. It thus follows that A3B—the only constitutively nuclear human A3 enzyme—may pose a threat to the genomic integrity of viruses that undergo replication in the nucleus, such as EBV and related herpesviruses.

Viruses that are susceptible to restriction by cellular A3 enzymes have evolved equally potent counteraction mechanisms to ensure successful replication (reviewed by [2, 3, 19]). The most well-studied A3 antagonist is the lentiviral accessory protein Vif, which nucleates the formation of an E3 ubiquitin ligase complex to promote the selective degradation of up to five different cytoplasmic A3 enzymes: A3C, A3D, A3F, A3G, and A3H (reviewed by [2, 20, 21]).

Extensive mutagenesis analyses revealed that HIV-1 Vif evolved to use distinct binding surfaces in order to recognize different A3s, as underscored by separation-of-function mutants of Vif that selectively degrade only a subset of these enzymes (reviewed by [20, 22, 23]). These studies have been essential to our understanding of how viruses adapt to overcome innate host defenses and, collectively, highlight the importance of elucidating structure-function relationships between proteins at the host-pathogen interface.

HIV-1 and related retroviruses are thought to be primary targets of A3 enzymes due to obligate ssDNA replication intermediates, abundant evidence for A3-catalyzed hypermutation of viral genomes, and a potent A3 counteraction mechanism. Reports of DNA viruses exhibiting signatures of A3-mediated mutagenesis suggest that these viruses may also be susceptible to restriction by A3 enzymes [2, 24], although the identification of viral antagonists has been challenging. We recently discovered that the large double-stranded (ds)DNA herpesvirus EBV uses a two-pronged approach to counteract the antiviral activity of A3B. The large subunit of the EBV ribonucleotide reductase (RNR), BORF2, directly inhibits A3B catalytic activity and relocalizes this enzyme from the nucleus to cytoplasmic aggregates—most likely to protect viral genomes from deamination during lytic replication [15]. Consistent with this interpretation, BORF2-null EBV accumulates A3B signature C/G to T/A mutations upon lytic reactivation and exhibits significantly lower infectivity [15].

The deamination activity of all A3 family enzymes requires a zinc-coordinating motif, which forms the core of the catalytic site and is surrounded by three loop regions (L1, L3, and L7) that contribute to binding of ssDNA substrates [25-27]. These loops vary in size and amino acid composition and account for differences in A3 catalytic rates and local target sequence preferences (reviewed by [28, 29]). The neutralization of A3B by EBV BORF2 is the first example of direct

inhibition of A3 ssDNA deaminase activity by another protein, and our recently published cryo-EM structure of EBV BORF2 in complex with the catalytic domain of A3B (A3Bctd) provides critical insights into the structural basis and molecular mechanism of this unanticipated host-pathogen interaction [30]. Specifically, our cryo-EM studies revealed that EBV BORF2 binds to the L1 and L7 regions of A3Bctd, resulting in a high-affinity interaction that effectively blocks the A3B active site from binding to ssDNA substrates and catalyzing deamination. Recently, we and others have shown that, in addition to EBV BORF2, the large RNR subunit from related human herpesviruses can also bind and relocalize A3B—as well as the highly similar A3A enzyme—suggesting that this A3 counteraction mechanism may be conserved evolutionarily [31-33].

Here, we investigate the molecular and functional conservation of A3-antagonism by the RNR subunits of different herpesviruses using structurally and evolutionarily informed approaches. We find that the RNR subunits of EBV and the related human gamma-herpesvirus Kaposi's sarcoma-associated herpesvirus (KSHV) interact with A3B and A3A by binding to distinct loop regions near the catalytic site, which accounts for differences in binding selectivity and capacity to inhibit A3 enzymatic activity. We also uncover potential coevolutionary interactions between A3 enzymes and primate gamma-herpesviruses by discovering a link between the generation of the *A3B* gene in primates and the evolution of A3-neutralization functionalities in viral RNRs. Last, we reconstruct an ancestral primate A3B enzyme and show that it is bound, relocalized, and inhibited by the RNRs of gamma-herpesviruses that infect primate species that encode *A3B* but not by the RNRs from primate viruses that infect species that naturally lack this gene. Taken together, our data provide evidence of a conserved and likely ancient A3 antagonism mechanism in primate gamma-herpesviruses and offer insights into how genetic conflicts with host

restriction factors may have affected—and likely continue to affect—the structure and function of viral proteins such as the herpesvirus RNR.

Results

Structural differences and similarities between EBV and KSHV RNR large subunits

Our recent cryo-EM studies of EBV BORF2-A3B complexes revealed a large binding surface composed of multiple structural elements from each protein [30]. Specifically, loops 1 (L1) and 7 (L7) of the A3B catalytic domain (A3Bctd) are sequestered by multiple interactions with BORF2 including contacts with residues L133 and Y134 of a novel short helix insertion (SHI) and residues R484 and Y481 of an additional helical loop structure (HLS). Y134 additionally forms a network of stabilizing interactions with other BORF2 residues including Q476 and R484 ([30]; depicted in **Figure 1A**). The BORF2 homolog from related gamma-herpesvirus KSHV (ORF61) has also been shown to bind A3B [15, 32]. However, an amino acid sequence alignment of EBV BORF2 and KSHV ORF61 indicates limited identity in the SHI and HLS regions (**Figure 1B**; complete sequence alignment in **Figure 1—figure supplement 1A**). We therefore used RoseTTAFold [34] to generate a structural model of KSHV ORF61 for 3-dimensional comparisons (**Figure 1C**). Although sharing only ~42% primary amino acid identity, the predicted model of KSHV ORF61 overlays well with the cryo-EM structure of EBV BORF2 with root mean square deviation (rmsd) of 1.26 Å. Interestingly, in addition to similar HLS regions, EBV BORF2 and the predicted structure of KSHV ORF61 share a similar sized SHI region (depicted in red in **Figure 1C**), which is lacking in nonviral class Ia RNRs such as those from *E. coli* and humans [30].

To gain further insights into the potential role of KSHV ORF61 SHI and HLS regions in binding to A3B, we used *in silico* docking to investigate how ORF61 and A3Bctd might interact. RoseTTAFold models of KSHV ORF61 and A3Bctd were aligned to the cryo-EM BORF2-A3Bctd complex using PyMOL [35] to create an initial docking pose followed by local protein-protein docking using RosettaDock 4.0 [36, 37]. Models were ranked by CAPRI-score and the best structure was selected based on overall interface energy score (**Methods** and **Figure 1—figure supplement 1B**). Several high-scoring models suggested that the ORF61 SHI region may interact with the loop 3 (L3) region of A3B, instead of L1 and L7 as for BORF2 (**Figure 1D-E**). Additionally, the predicted HLS region of KSHV ORF61 appears to be positioned further away from the RNR-A3Bctd binding interface and may not contribute to the interaction (**Figure 1E**). Taken together, these *in silico* predictions suggest that EBV BORF2 and KSHV ORF61 may differentially engage the catalytic domain of A3B by interacting with distinct loop regions. This possibility is tested in the following sections.

Involvement of the A3B and A3A loop 7 regions in binding to gamma-herpesvirus RNRs

Recent biochemical and structural studies demonstrated that direct contacts between EBV BORF2 and the L7 region of A3Bctd are essential for this host-pathogen interaction [15, 30]. A3Bctd is closely related to A3A (>90% amino acid identity including identical L7 regions; **Figure 2A**), suggesting that a similar interaction with L7 could explain EBV BORF2 ability to engage A3B and A3A, but not other A3 family members. Recent work also revealed that KSHV ORF61 is similarly capable of binding and relocalizing A3B and A3A, but not other A3 enzymes [32].

To test whether binding to L7 is a conserved requirement for viral RNR-A3 interaction, we generated a series of chimeric A3 proteins with reciprocal L7 swaps and tested their ability to bind to EBV BORF2 and KSHV ORF61 in co-immunoprecipitation (co-IP) experiments (**Figure 2B**). 293T cells were transfected with FLAG-tagged viral RNRs and EGFP-tagged A3s, subjected to anti-FLAG affinity purification, and analyzed by immunoblotting. As expected from our prior studies, replacing L7 of A3Bctd or A3A with L7 of A3Gctd completely abolishes interaction with EBV BORF2 (**Figure 2B**, left panel, lanes 1-4) and, reciprocally, replacing L7 of A3Gctd with L7 of A3A/Bctd endows binding (**Figure 2B**, left panel, lanes 5-6). However, we were surprised to find that replacing L7 of A3Bctd or A3A with L7 of A3Gctd does not affect binding by KSHV ORF61 (**Figure 2B**, right panel, lanes 1-4). Accordingly, replacing L7 of A3Gctd with L7 of A3A/Bctd does not enable pull-down by KSHV ORF61 (**Figure 2B**, right panel, lanes 5-6)

An additional defining feature of the interaction between A3B/A and viral RNRs is relocalization from the nucleus to the cytoplasm [15, 31, 32]. To evaluate the subcellular localization phenotypes of the L7 chimeras described above, we performed quantitative fluorescence microscopy experiments using HeLa cells co-transfected with EGFP-tagged A3s and FLAG-tagged viral RNRs (**Figure 2C**, representative images). In agreement with our co-IP results, replacing L7 of A3Bctd or A3A with L7 of A3Gctd markedly decreases EBV BORF2-mediated relocalization, although low levels of A3Bctd L7G relocalization can still be detected in a subset of cells. Reciprocally, replacing L7 of A3Gctd with L7 of A3A/Bctd enables effective relocalization (**Figure 2C**, representative images; **Figure 2D-E**, quantification). In further agreement with our co-IP results, swapping L7 does not decrease KSHV ORF61-mediated relocalization of A3Bctd or A3A, nor does it increase A3Gctd relocalization (**Figure 2C**, representative images; **Figure 2D-E**, quantification). Collectively, our experimental data support

the *in silico* protein docking predictions and indicate that, despite an overall high degree of predicted structural similarity, the RNR subunits from EBV and KSHV engage A3Bctd and A3A through at least partially distinct surfaces.

Role of A3B and A3A loop 3 in binding to gamma-herpesvirus RNRs

L3 is one of three loops that surround the A3 active site (**Figure 2A**). In our recent cryo-EM studies of the EBV BORF2-A3Bctd complex, we were unable to model A3Bctd L3 residues due to the flexibility of this region ([30]; dashed region in **Figure 3A**). *In silico* protein-protein docking predictions in **Figure 1D-E** suggest that this loop may be important for KSHV ORF61 binding. Therefore, to examine whether L3 participates in viral RNR-A3 interactions, we performed co-IP experiments using a panel of chimeric A3 proteins with reciprocal L3 swaps (**Figure 3B**).

Once again, we observed markedly distinct phenotypes in co-IPs with EBV BORF2 and KSHV ORF61. For EBV BORF2, similar pull-down levels are observed for WT versus L3-swapped A3s (**Figure 3B**, left panel, lanes 1-4), whereas replacing L3 of A3Bctd or A3A with L3 of A3Gctd nearly abolishes the interaction with KSHV ORF61 (**Figure 3B**, right panel, lanes 1-4). Additionally, replacing L3 of A3Gctd with L3 of A3A/Bctd is sufficient to promote modest immunoprecipitation with KSHV ORF61 (**Figure 3B**, right panel, lanes 5-6). These results indicate that the L3 region is critical for KSHV ORF61 binding and likely to have little or no role in binding to EBV BORF2.

To evaluate the subcellular localization phenotypes of the L3 chimeras described above, we performed quantitative fluorescent microscopy experiments using HeLa cells (**Figure 3C**, representative images). Similar to our L7 results, microscopy observations are largely concordant

with co-IP phenotypes. For EBV BORF2, L3 swaps have no substantial effect on A3 relocalization, whereas replacing L3 of A3Bctd or A3A with the corresponding region of A3Gctd causes a marked decrease in relocalization by KSHV ORF61 (**Figure 3C**, representative images; **Figure 3D-E**, quantification). Notably, KSHV ORF61 is able to relocalize the chimeric A3Gctd L3A/Bctd protein, albeit to considerably lesser extents compared to WT A3Bctd or A3A. These data provide further support for the *in silico* predictions above by demonstrating the importance of the L3 region for binding to KSHV ORF61 and promoting cytoplasmic aggregate formation in living cells.

Role of A3B and A3A loop 1 in binding to gamma-herpesvirus RNRs

An unanticipated discovery made possible by our recent EBV BORF2-A3Bctd cryo-EM structure is the existence of an extensive interaction between the short HLS region of BORF2 and L1 of A3Bctd [30]. This L1 interaction accounts for EBV BORF2 binding preferentially to A3B as opposed to A3A. We therefore compared the role of the L1 region of A3Bctd versus A3A in the interaction of these enzymes with EBV BORF2 and KSHV ORF61. First, we confirmed our recent report by showing that replacing L1 of A3A with L1 of A3Bctd leads to a marked increase in EBV BORF2 co-IP (**Figure 4A**, left panel, lanes 1-2). A strong increase in pull-down is also visible when L1 of A3A is replaced with L1 of A3Gctd, possibly due to the similarities between the L1 regions of A3Bctd and A3Gctd, where each enzyme has a 3-residue motif that is absent in A3A (PLV and PWV, respectively; **Figure 2A**). Reciprocally, replacing L1 of A3Bctd with L1 of A3A—but not with L1 of A3Gctd—dramatically diminishes the interaction with EBV BORF2 (**Figure 4A**, left panel, lanes 4-5).

In contrast, KSHV ORF61 does not appear to bind more strongly to A3Bctd compared to

A3A (**Figure 4A**, right panel, compare lanes 1 and 4). Moreover, L1 swaps between A3Bctd and A3A have no effect on KSHV ORF61 co-IPs (**Figure 4A**, right panel, lanes 1, 2, 4, and 5). However, interestingly, replacing L1 of either A3Bctd or A3A with residues from A3Gctd decreases interaction with KSHV ORF61 (**Figure 4A**, right panel, lanes 3 and 6). This may be attributable to the presence of a unique bulky tryptophan residue in A3Gctd L1 potentially clashing with a nearby interacting surface (**Figure 1** and **Figure 2A**).

We next assessed the subcellular localization phenotypes of these L1 chimeras using quantitative fluorescence microscopy (**Figure 4B**, representative images; **Figure 4C-D**, quantification). In contrast to the co-IP results above, swapping the L1 regions of A3Bctd and A3A does not overtly affect relocalization by EBV BORF2. This agrees with our published results [30] and supports a model in which compensatory avidity interactions during cytoplasmic aggregate formation are likely to account for the similar relocalization phenotypes of A3Bctd and A3A despite BORF2 interacting more weakly with the latter protein. In comparison, for KSHV ORF61, swapping L1 of A3Bctd with L1 of A3Gctd slightly decreases relocalization, whereas the remaining L1 swaps have no significant effect (**Figure 4B**, representative images; **Figure 4C-D**, quantification).

EBV BORF2 interacts with A3Bctd loop 1 to inhibit DNA deaminase activity

In addition to causing protein relocalization, direct binding by EBV BORF2 potently inhibits A3B catalytic activity *in vitro* [15, 30]. This supports a model in which EBV BORF2 simultaneously inhibits A3B activity and moves the complex away from viral replication centers to protect EBV genomes from deamination during lytic reactivation. Given results here showing

that EBV BORF2 and KSHV ORF61 bind to both A3B and A3A, we next asked whether direct enzymatic inhibition is a conserved feature of this pathogen-host interaction.

HeLa T-REx cells stably expressing doxycycline-inducible A3B-EGFP or A3A-EGFP were transfected with an empty vector control or FLAG-tagged EBV BORF2 and KSHV ORF61. After 24 hours of doxycycline induction, whole cell extracts were incubated with a fluorescently labelled ssDNA oligonucleotide substrate containing a single A3B/A deamination motif (5'-TC). A3B/A-mediated deamination of cytosine-to-uracil followed by uracil excision by UDG and abasic site cleavage by sodium hydroxide (NaOH) yields a shorter product that can be detected by SDS-PAGE and fluorescence scanning (see **Methods**). As expected, EBV BORF2 strongly inhibits A3B deaminase activity in cell lysates as indicated by minimal product accumulation (**Figure 5A**, left panel, compare lanes 1 and 2). KSHV ORF61 also inhibits A3B-catalyzed deamination when compared to a no-RNR control, albeit to a considerably lesser extent compared to EBV BORF2 (**Figure 5A**, left panel, compare lanes 1 and 3).

Surprisingly, neither EBV BORF2 nor KSHV ORF61 substantially inhibits the deaminase activity of A3A (**Figure 5A**, right panel). This stark difference is supported by our cryo-EM structure showing extensive contacts between BORF2 and the L1 and L7 regions of A3Bctd, which results in a high-affinity interaction that blocks the A3B active site from binding to ssDNA substrates and catalyzing deamination [30]. Specifically, A3B L1 residues 206-PLV-208 contribute to the formation of a hydrophobic pocket in which EBV BORF2 residue Y481 is inserted (depicted in **Figure 5B**). In contrast, this 3-residue motif is absent in A3A L1 (**Figure 2A**). We therefore hypothesized that inhibition of deaminase activity by EBV BORF2 requires not only binding to the catalytic domain through L7, but also a significant interaction with L1.

To further test this idea, we first generated mutants of EBV BORF2 Y481 and examined their ability to inhibit A3B-catalyzed deamination using the HeLa T-REx A3B-EGFP system described above. Mutating EBV BORF2 Y481 to an alanine or a proline effectively prevents inhibition of A3B-catalyzed deamination, likely by weakening the interaction with the L1 region (**Figure 5C**). Next, we compared the ability of EBV BORF2 to inhibit the deaminase activity of WT versus L1 swapped A3B and A3A in HeLa cells transfected with FLAG-tagged BORF2 and EGFP-tagged A3s. In agreement with results from **Figure 5A**, WT A3B deaminase activity is potently inhibited by EBV BORF2 (**Figure 5D**, left, lanes 1 and 2). In contrast, when L1 of A3B is replaced with L1 of A3A, EBV BORF2 completely loses its capacity to inhibit ssDNA deamination (**Figure 5D**, left, lanes 3 and 4). In further agreement with data in **Figure 5A**, WT A3A deaminase activity is not inhibited by EBV BORF2 (**Figure 5D**, right, lanes 1 and 2). Unfortunately, the chimeric A3A L1B construct exhibits very low levels of ssDNA deaminase activity, and it is therefore not possible to determine the susceptibility to inhibition by EBV BORF2 (**Figure 5D**, right, lanes 3 and 4). Nevertheless, these results combine to indicate that the EBV BORF2 interaction with A3B L1 is a requirement for inhibiting deamination and also help explain why A3A activity is not inhibited similarly.

A structural overlay of the cryo-EM structure of the EBV BORF2-A3B complex and the RoseTTAFold-predicted model of KSHV ORF61 indicates that a residue analogous to the interlocking Y481 of BORF2 is lacking in ORF61 (**Figure 5E**). Instead, KSHV ORF61 has a proline (P488) immediately downstream of shared proline and threonine residues. This difference suggests that L1 may be allowed to remain flexible and engage ssDNA substrates when A3B is bound to KSHV ORF61. Alternatively—and perhaps more likely—a lack of interaction with L1 may result in RNR-A3B complexes with lower binding affinity (i.e., lower association rate and/or

higher dissociation rate), thereby rendering KSHV ORF61 unable to inhibit A3B-catalyzed deamination to the same extent as EBV BORF2.

The genesis of *A3B* may have shaped gamma-herpesvirus RNR evolution in primates

A3 genes are unique to placental mammals and the evolutionary history of the *A3* locus is marked by frequent expansions and contractions resulting from independent events of gene duplication, loss, and fusion [38-41]. For instance, in the primate lineage leading to humans the *A3* locus evolved from an ancestral 3 deaminase domain cluster to the present day 11 domain (7 gene) cluster. Interestingly, phylogenetic and comparative genomics studies indicate that *A3B* may be the youngest *A3* family member in primates [39, 42]. Specifically, *A3B* was likely generated the by duplication of ancestral deaminase domains in the common ancestor of Catarrhini primates (i.e., hominoids and Old World monkeys) and is consequently absent in Platyrrhini primates (i.e., New World monkeys) (depicted in **Figure 6A**). We therefore hypothesized that the function of an ancestral gamma-herpesvirus RNR as an innate immune antagonist may have been selected by the birth of *A3B* during this critical period of primate evolution (i.e., 43 to 29 million years ago, after the split of Catarrhini and Platyrrhini primates).

To test this idea, we cloned the RNR large subunit genes from representative gamma-herpesviruses that infect primates with A3B (EBV and KSHV, which infect humans, and McHV-4 and McHV-5, which infect rhesus macaques) and related viruses that infect primates without A3B (CalHV-3 and SaHV-2, which infect New World marmosets and squirrel monkeys, respectively) (**Figure 6B**). Of note, these RNRs represent gamma-herpesviruses from two evolutionarily diverged genera—*Lymphocryptovirus* and *Rhadinovirus* (**Figure 6B**). We then examined the capacity of these viral RNRs to bind to human A3Bctd and A3A in co-IP experiments

(**Figure 6C**). In support of our hypothesis, only the RNRs from viruses that infect A3B-encoding primates (humans and rhesus macaques) pull-down human A3Bctd and A3A (**Figure 6C**, lanes 1-4, left and right panels). In contrast, the RNRs from viruses that infect primates without A3B (marmosets and squirrel monkeys) are unable to pull-down either enzyme (**Figure 6C**, lanes 5-6, left and right panels).

We next asked if these viral RNRs can relocalize human A3B and A3A. HeLa T-REx cells were transfected with FLAG-tagged viral RNRs and incubated with doxycycline to induce A3B-EGFP or A3A-EGFP expression (**Figure 6D**). In agreement with our co-IP results, only the RNRs from viruses that infect A3B-encoding primates are able to relocalize human A3B and A3A from the nucleus to the cytoplasmic compartment. Together, these results suggest that the ability to bind and relocalize human A3 enzymes is conserved among RNRs from gamma-herpesviruses that infect Catarrhini primates.

We additionally used HeLa T-Rex cells stably expressing doxycycline-inducible human A3B-EGFP or human A3A-EGFP to ask whether these viral RNRs are capable of inhibiting the ssDNA deaminase activity of human A3B and A3A (**Figure 6E**). Consistent with our previous results, EBV BORF2 potently inhibits A3B (**Figure 6E**, top panel, lane 2). A similarly strong inhibition is seen with the RNR of rhesus macaque *Lymphocryptovirus* McHV-4 (**Figure 6E**, top panel, lane 3). The RNRs from KSHV and rhesus macaque *Rhadinovirus* McHV-5 did not overtly affect the formation of a product band but appeared to slow down A3B-catalyzed deamination, as suggested by a stronger substrate band compared to the no-RNR control (**Figure 6E**, top panel, lanes 5 and 6). Most importantly, the RNRs from marmoset *Lymphocryptovirus* CalHV-3 and squirrel monkey *Rhadinovirus* SaHV-2 are both defective in inhibiting A3B activity (**Figure 6E**, top panel, lanes 4 and 7). Additionally, none of the six different viral RNRs inhibits the activity of

human A3A (**Figure 6E**, bottom panel). Collectively, our results suggest a possible link between the birth of *A3B* in the common ancestor of Catarrhini primates and the evolution of A3B-counteraction functionalities in an ancestral gamma-herpesvirus RNR large subunit (i.e., direct binding, subcellular relocalization, and enzymatic inhibition).

An short helical loop structure from EBV BORF2 enables the marmoset CalHV-3 RNR to bind to human A3B and A3A

A3 enzymes are composed of single or double zinc-coordinating (Z) domains that can be grouped into three classes (Z1, Z2, and Z3) based on sequence similarity [43]. In humans, A3A, A3C, and A3H are single Z domain proteins (Z1, Z2, and Z3, respectively), whereas A3B, A3G, A3D, and A3F harbor two Z domains (Z2-Z1 for A3B and A3G, Z2-Z2 for A3D and A3F). Although New World monkeys lack A3B, the genomes of these species—as well as all primates—encode a single-domain Z1 enzyme homologous to human A3A. Therefore, it is possible that the RNRs of gamma-herpesviruses that infect New World monkeys may have evolved to antagonize the A3A enzymes of their natural host species.

To investigate this possibility, we performed co-IP experiments using FLAG-tagged viral RNRs and host-matched EGFP-tagged A3As (**Figure 7A**). In line with our results above with human A3Bctd and A3A (**Figure 6**), all four RNRs from viruses that infect Catarrhini primates bind to the A3A enzymes of their respective host species (**Figure 7A**, lanes 1, 2, 4, and 5). In contrast, the RNRs from viruses infecting New World monkeys are unable to bind to their hosts' A3As (**Figure 7A**, lanes 3 and 6). We next asked if these RNRs can relocalize the A3A enzymes encoded by their respective host species (**Figure 7B**). In agreement with the co-IP results, the RNRs from EBV/KSHV and McHV-4/McHV-5 relocalize human and rhesus A3A, respectively

(**Figure 7B**, left and center). In contrast, the RNRs from CalHV-3 and SaHV-2 do not alter the localization of marmoset and squirrel monkey A3A, respectively (**Figure 7B**, right). These results further support the idea that A3 antagonism by gamma-herpesvirus RNRs is likely to be limited to viruses that infect primates with *A3B* and, additionally, suggest that binding to A3A may be a consequence of similarity to A3Bctd (identical loop 3 and loop 7 regions and >90% overall identity).

Our recent cryo-EM structure of the EBV BORF2-A3Bctd complex revealed that loops 1 and 7 of A3B are sequestered by multiple interactions with BORF2 including contacts with a short helical loop structure (HLS) involving the critical BORF2 residues Y481 and R484 ([30]; region depicted in red in **Figure 7C** and alternative poses in **Figure 1** and **Figure 5**). Amino acid sequence alignments of the RNRs from human EBV, rhesus macaque McHV-4, and marmoset CalHV-3, which are all members of the *Lymphocryptovirus* genus, show that the former two RNRs share identical amino acid sequences in the HLS region, whereas the latter protein has multiple differences (HLS region alignment in **Figure 7C**). Given the overall similarity between the 3-dimensional architecture of EBV BORF2 and the predicted structure of CalHV-3 RNR (**Figure 7—figure supplement 1**), we next asked whether replacing the HLS region of the CalHV-3 RNR with the corresponding motif from EBV BORF2 may be sufficient to enable binding to A3B/A enzymes in co-IP experiments. Interestingly, although the chimeric CalHV-3 RNR construct is still unable to bind marmoset A3A (**Figure 7D**, lane 2), this protein is now able to efficiently bind to both human A3Bctd and A3A (**Figure 7D**, lanes 3 and 4). These results indicate that the HLS region of EBV BORF2 is sufficient to endow a normally non-binding *Lymphocryptovirus* RNR with a capacity to engage both human A3Bctd and A3A, likely by enabling a direct physical interaction with the identical loop 7 region of these enzymes (L7 region alignment in **Figure 7E**).

1

2 **Ancestral A3B is active and antagonized by present-day RNRs from gamma-herpesviruses**
3 **that infect *A3B*-encoding primates**

4 To gain deeper insights into the interaction between A3B and gamma-herpesvirus RNRs,
5 we generated a phylogenetic tree of all publicly available primate A3B sequences and used the
6 GRASP-suite [44] to reconstruct the amino acid sequence of ancestral A3B (**Figure 8A** and **Figure**
7 **8—figure supplement 1**). We then examined an expanded panel of primate gamma-herpesvirus
8 RNRs (i.e., 11 different viral proteins) for a capacity to bind human and ancestral A3Bctd in co-
9 IP experiments (**Figure 8B**). In agreement with results above, only the RNRs from gamma-
10 herpesviruses that infect primates with A3B are capable of binding to human A3Bctd (**Figure 8B**,
11 left). Similarly, nearly all RNRs from viruses that infect primates with A3B are able to bind the
12 ancestral protein and, importantly, all viruses that infect primates which naturally lack the *A3B*
13 gene are unable to bind the ancestral enzyme (**Figure 8B**, right).

14 To further explore the properties of the ancestral A3B protein, we performed a
15 comprehensive series of subcellular localization experiments and ssDNA deaminase activity
16 assays using this expanded panel of primate gamma-herpesvirus RNRs (**Figure 8C-D** and **Figure**
17 **8—figure supplement 2A**). As above, human A3B is completely relocalized from the nucleus to
18 the cytoplasmic compartment by the RNRs of viruses that infect humans and other primates with
19 an *A3B* gene (**Figure 8C**, top; **Figure 8—figure supplement 2A**). Similar results were obtained
20 with ancestral A3B, although only partial relocalization is observed with the KSHV ORF61
21 (**Figure 8C**, bottom; **Figure 8—figure supplement 2A**). KSHV ORF61also exhibit a decreased
22 ability to inhibit the deaminase activity of ancestral A3B compared to human A3B (**Figure 8D**,
23 compare lane 5 from the top and bottom panels). These relocalization and ssDNA deaminase

activity results correlate with the co-IP observations in **Figure 8B**, where the RNRs from the two human gamma-herpesviruses showed decreased binding to ancestral A3Bctd compared to the human enzyme. Such differences in binding to human versus ancestral A3B may be an indication of adaptive changes in present day viral RNRs driven by host-pathogen conflicts.

A phylogenetic tree based on the amino acid sequences of available gamma-herpesvirus RNRs indicates that the sequences from human, rhesus macaque, and marmoset *Lymphocryptoviruses* cluster together and the sequences from the different *Rhadinoviruses* cluster according to viral lineage (**Figure 8E**). Interestingly, despite its similarity to KSHV ORF61, the RNR from ColHV-1—a novel KSHV-related virus identified in African *Colobinae* monkeys [45]—differs from the RNRs of all other viruses that infect Catarrhini primates tested here in that is shows no ability to bind, relocalize, or inhibit present-day human or ancestral A3B (**Figure 8B**, **Figure 8—figure supplement 2A**, and **Figure 8D**) or interact with its host's A3A (**Figure 8—figure supplement 3A**). This apparent incongruity may be explained by the fact that African *Colobinae* monkeys appear to have lost the *A3B* gene following the *Colobinae* subfamily split into the African and Asian tribes 10 to 14 million years ago (**Figure 8—figure supplement 3B-C**), which presumably relieved the pressure for ColHV-1 to maintain an A3B-neutralization activity.

Differential binding of viral RNRs to present day and ancestral A3B is likely a reflection of past—and possibly still ongoing—genetic conflicts. To further explore this idea, we analyzed a data set of primate *A3B* DNA sequences for evidence of codon sites under diversifying/positive selection using the mixed effects model of evolution algorithm (MEME) [46]. We detected evidence for diversifying selection at multiple sites, including clusters of positively selected residues in loops 1 and 3 (**Figure 8F**, top; **Figure 8—figure supplement 4A-B**). This suggests that the amino acid composition of these regions may have been influenced by recurrent

diversifying selection, possibly as a mean to escape binding by gamma-herpesvirus RNRs, which in turn would select for virus adaptation to restore binding (depicted in **Figure 8F**, bottom). In line with this interpretation and together with results in **Figure 3** and **Figure 4**, it is possible that differences in positively selected residues in loops 1 and 3 may respectively account for the apparent increased ability of EBV BORF2 and KSHV ORF61 to antagonize present day human A3B compared to ancestral A3B.

Discussion

Viruses use a plethora of mechanisms to avoid and antagonize host immune responses. We previously reported an unexpected role for herpesvirus RNRs in counteracting innate immunity by revealing that EBV repurposes its RNR subunit BORF2 to neutralize the antiviral activity of cellular A3B [15]. The importance of this A3B counteraction mechanism is evidenced by BORF2-null EBV exhibiting lower infectivity and an accumulation of A3B signature C/G-to-T/A mutations [15]. Subsequent studies revealed that the RNR subunits from related human herpesviruses also bind and relocalize A3B, as well as the highly similar A3A enzyme [31, 32]. Here, we leverage knowledge gained from our recent cryo-EM structure of EBV BORF2 in complex with the catalytic domain of A3B [30] to investigate the mechanistic conservation of A3 antagonism by herpesviruses. We find that EBV BORF2 and KSHV ORF61 bind to cellular A3 enzymes via distinct surfaces and show that direct inhibition of ssDNA deaminase activity by EBV BORF2 requires specific interactions with the L1 of A3B in addition to the L7 region. We further find that across different primate gamma-herpesvirus lineages, the existence of an RNR-mediated mechanism to counteract cellular A3 enzymes appears to be confined to viral species that infect A3B-expressing primates. Together, these results suggest that the genesis of *A3B* by duplication

of ancestral Z2 and Z1 deaminase domains in the common ancestor of present-day hominoids and Old World monkeys may have selected for an ancestral gamma-herpesviruses with a capacity to antagonize this new restriction factor through its large RNR subunit. To our knowledge, this is the first study to link the birth of an antiviral host factor and the evolution of an equally potent viral escape mechanism.

Our data here highlight the overall conservation of A3B antagonism by gamma-herpesvirus large RNR subunits but also reveal new and unexpected molecular differences. In agreement with our recent cryo-EM studies [30], we show that A3B L7 is a critical determinant of EBV BORF2 binding, whereas A3B L1 serves to help strengthen this interaction and confer selective binding over A3A (**Figure 2** and **Figure 4**). Surprisingly, these same rules do not apply to KSHV ORF61, which appears to mainly rely on L3 to bind and relocalize cellular A3B/A (**Figure 3**). Together, our results suggest a model where interactions with L7 and/or L3—which are identical in A3B and A3A, but not in any other members of the A3 family—enable gamma-herpesvirus RNRs to bind and relocalize both enzymes, whereas additional interacting regions (i.e., EBV BORF2 HLS and A3B L1) may help confer selectivity and additional neutralization activity (e.g., potent inhibition of A3B ssDNA deaminase activity). It is additionally interesting that this may be a new manifestation of a pathogen-host wobble mechanism [47] in which ongoing small changes over evolutionary time in the viral RNRs can manifest molecularly in the present-day target protein as distinct binding surfaces (e.g., EBV BORF2 requires A3Bctd L7 and KSHV ORF61 requires L3).

Using a panel of 11 primate gamma-herpesvirus RNRs, we show that, in addition to binding and relocalization, all RNRs derived from viruses that infect hosts that encode A3B show some capacity to inhibit the ssDNA deamination activity of human A3B, but none of these viral proteins could inhibit the deaminase activity of human A3A (**Figure 8** and **Figure 8—figure**

supplement 2B). This is consistent with a model in which the antiviral activity of A3B may have forced these viruses to evolve and maintain a potent counteraction mechanism, whereas interaction with A3A may be a consequence of sequence homology to A3B and less likely a true physiologic function. This interpretation is further supported by our observation that the RNRs from viruses that infect *A3B*-null primates are incapable of interacting with the A3A enzymes of their natural host species (**Figure 7**).

The A3B-selective nature of this host-pathogen interaction is orthogonally supported by previous binding studies showing that the EBV BORF2-A3Bctd interaction has a biolayer interferometry dissociation constant of ~1 nM, whereas the interaction with the >90% identical A3A enzyme is over 10-fold weaker [30]. Our results here reveal that, in addition to providing selective binding to A3B, EBV BORF2 interaction with the L1 region of A3B is required for potently inhibiting DNA deaminase activity (**Figure 5**). It is thus possible that the direct binding of EBV BORF2 to two loop regions near the A3B catalytic site—L1 and L7—helps stabilize the BORF2-A3B complex, resulting in a high affinity interaction and consequently blocking the A3B active site from accessing ssDNA substrates and catalyzing deamination. Importantly, the potential adaptability of this A3B counteraction mechanism in *Lymphocryptovirus*es is highlighted by results showing that grafting the HLS region from EBV BORF2—which contains residues that directly contact A3B L1 and L7—into the RNR from marmoset CalHV-3 enables binding to human A3B (**Figure 7**). In comparison, KSHV ORF61 binding to A3B largely through L3 may result in a lower affinity interaction that more weakly inhibits ssDNA deamination. These inferences support a model in which binding to the A3 active site by viral RNRs is responsible for both protein relocalization and enzymatic inhibition, yet substantial deaminase activity inhibition is only observed above a certain affinity threshold. We therefore postulate that preferential binding

to A3B might have been an evolutionary adaptation resulting from selective pressures disproportionally affecting gamma-herpesviruses within the *Lymphocryptovirus* genus, such as EBV and McHV-4. For instance, EBV latency is limited to B lymphocytes [48], which can express high levels of endogenous A3B [49, 50], strongly implying that the neutralization of this enzyme is required to preserve viral genome integrity upon lytic reactivation. On the other hand, the broader tropism of KSHV suggests that lytic replication might not always occur in the presence of A3B [51-55], which may result in lower exposure to A3B-mediated deamination.

Evolutionary analyses of host-pathogen antagonistic relationships are powerful tools for understanding protein functions. The link between *A3B* generation by duplication of ancestral Z2 and Z1 domains and the evolution of A3-neutralization functionalities in herpesvirus RNRs described here suggests that the double domain architecture of A3B may be a requirement for herpesvirus restriction. Notably, the non-catalytic Z2 domain of A3B has been shown to account for this enzyme's distinct nuclear localization [56] and association with multiple RNA binding proteins, in addition to allosterically regulating the catalytic activity of the Z1 domain [57]. It is therefore plausible that cooperation between Z2 and Z1 domains accounts for A3B antiviral function, possibly by facilitating access to herpesvirus transcription complexes inside the nucleus (e.g., ssDNA exposed in viral R-loop structures [58]). Future studies dissecting the molecular mechanisms that allow A3B to recognize and target foreign nucleic acids such as herpesvirus genomes might reveal insights into the regulation of ssDNA deaminase activity and how the loss of such regulatory mechanisms can at times result in cellular chromosomal DNA mutations and cancer development (reviewed by [59, 60]). Future work will likely benefit from interrogating the RNRs and A3 enzymes from a larger and more diverse panel of virus and host species, which is currently prevented by the relatively small number of herpesvirus complete genomes deposited in

public databases. Bat species may be particularly interesting given they encode the largest known A3 repertoire in mammals [61, 62] and are also known to host a variety of different herpesviruses [63-73].

Materials and methods

DNA constructs

The large RNR subunit from EBV (HHV-4), McHV-4, CalHV-3, KSHV (HHV-8), McHV-5, McHV-11, McHV-12, AtHV-3, SaHV-2, ColHV-1, and SgHV-1 were PCR amplified from gBlocks purchased from Integrated DNA Technologies and cloned into pcDNA4/TO (Invitrogen #V102020) with a C-terminal 3×FLAG-tag (DYKDDDDK). EBV BORF2$_{Y481A}$, EBV BORF2$_{Y481P}$, CalHV-3 RNR$_{WEARDV490YDSRDM}$ constructs were generated by site-directed mutagenesis. Wild-type constructs match the following GenBank accessions: EBV BORF2 (V01555.2), McHV-4 BORF2 (AY037858.1), CalHV-3 ORF55 (AF319782.2), KSHV ORF61 (U75698.1), McHV-5 ORF61 (AF083501.3), McHV-11 ORF61 (AY528864.1), McHV-12 ORF61 (KP265674.2), AtHV-3 ORF61 (AF083424.1), SaHV-2 ORF61 (X64346.1), ColHV-1 ORF61 (MH932584.1), and SgHV-1 ORF26 (OK337614.1).

APOBEC3 enzymes were cloned into pcDNA5/TO (Invitrogen #V103320) with a C-terminal EGFP-tag, as described previously [74]. The loop swap constructs A3B_L1A3A$_{25-30}$, A3Bctd$_{193-382}$_L1A3A$_{25-30}$, A3Bctd$_{193-382}$_L1A3G$_{209-217}$, A3Bctd$_{193-382}$_L3A3G$_{247-254}$, A3Bctd$_{193-382}$_L7A3G$_{317-322}$, A3A_L1A3B$_{205-213}$, A3A_L1A3G$_{209-217}$, A3A_ L3A3G$_{247-254}$, A3A_L7A3G$_{317-322}$, A3Gctd$_{197-384}$_L3A3A$_{60-67}$, and A3Gctd$_{197-384}$_L7A3A$_{132-137}$ were generated by site-directed mutagenesis. Wild-type constructs match the following GenBank accessions: *Homo sapiens* A3B

(NM_004900), *Homo sapiens* A3A (NM_145699), *Homo sapiens* A3G (NM_021822), *Macaca mulatta* A3A (NM_001246231.1), *Macaca mulatta* A3B (NM_001246230.2), *Callithrix jacchus* A3A (NM_001301845.1), *Saimiri boliviensis* A3A (XM_039462978.1), *Colobus angolensis palliatus* A3A (XM_011964212.1).

Q5 High-Fidelity DNA Polymerase (NEB #M0491) was used for site-directed mutagenesis and Phusion High-Fidelity DNA Polymerase (ThermoFisher #F530L) was used for all other PCR reactions. All constructs were confirmed by Oxford Nanopore whole plasmid sequencing.

Human cell culture

Unless indicated, cell lines were derived from established laboratory collections. All cell cultures were supplemented with 10% fetal bovine serum (Gibco #26140-079), 1× penicillin-streptomycin (Gibco #15140-122), and periodically tested for mycoplasma contamination using MycoAlert PLUS Mycoplasma Detection Kit (Lonza #LT07-710). 293T cells were cultured in high glucose DMEM (HyClone #SH30022.01) and HeLa cells were cultured in RPMI 1640 (Gibco #11875-093)

To generate HeLa T-REx A3A-EGFP and A3B-EGFP cell lines, HeLa cells were transduced with pLENTI6/TR (Invitrogen #V480-20), selected with 5 μg/ml blasticidin (GoldBio #B-800-25), and subcloned by serial dilution to obtain single cell clones. Clonal populations were stably transfected with pcDNA5/TO-A3A-EGFP or pcDNA5/TO-A3B-EGFP, selected with 200 μg/ml hygromycin (GoldBio #H-270-1), and subcloned by serial dilution. Representative clones were used for DNA deaminase activity assays (**Figure 5A**, **Figure 5C**, **Figure 6E**) and IF microscopy experiments (**Figure 6D**).

Co-immunoprecipitation experiments

Approximately 250,000 293T cells were seeded in 6-well plates and transfected the next day with 100 ng pcDNA5/TO-A3-EGFP and 200 ng pcDNA4/TO-RNR-3×FLAG using TransIT-LT1 (Mirus #2304). Cells were harvested ~24 hours after transfection and resuspended in 500 µl of lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 0.5% IGEPAL, 100 µg/ml RNaseA, 1 tablet of Sigma-Aldrich cOmplete Protease Inhibitor Cocktail). Resuspensions were vortexed, sonicated on ice-bath for 10 seconds at the lowest setting using a Branson sonifier, and centrifuged at 15,000 × g at 4 °C for 15 minutes. A 50 µl aliquot was removed from each sample for input detection and the remaining clarified whole cell lysate was incubated with 20 µl packed gel volume of anti-FLAG M2 magnetic beads (Sigma-Aldrich #M8823) at 4 °C overnight with gentle rotation. Samples were washed 3 times with 500 µL of lysis buffer and bound protein was eluted with 50 µl of elution buffer (100 µg/mL FLAG peptide (Sigma-Aldrich #F3290), 50 mM Tris-HCl pH 7.4, 150 mM NaCl). Input and elution samples were analyzed by immunoblot using mouse anti-GFP (1:5,000, Clontech #632381) to detect A3-EGFP, rabbit anti-FLAG (1:5,000, Sigma-Aldrich #F7425) to detect RNR-3×FLAG, and rabbit anti-GAPDH (1:5,000, Proteintech #60004) to detect GAPDH (loading control). Secondary antibodies used were IRDye 800CW anti-rabbit (LI-COR #926-32211) and HRP-linked anti-mouse (Cell Signaling #7076), both at 1:10,000 dilution.

Immunofluorescence microscopy experiments

In **Figure 6D**, approximately 5,000 engineered HeLa T-REx cells (described above) were seeded into 96-well imaging plates (Corning #3904) and transfected the next day with 30 ng pcDNA4/TO-RNR-3×FLAG. A3B-EGFP or A3A-EGFP expression was induced 6 hours after

transfection with 50 ng/mL doxycycline. Approximately 24 hours after induction, cells were fixed in 4% formaldehyde (ThermoFisher #28906), permeabilized in 0.2% Triton X-100 (Sigma-Aldrich #T8787) for 10 minutes and incubated in blocking buffer (2.8 mM KH2PO4, 7.2mM K2HPO4, 5% goat serum (Gibco #16210064), 5% glycerol, 1% cold water fish gelatin (Sigma- Aldrich #G7041), 0.04% sodium azide, pH 7.2) for 1 hour. Cells were then incubated with primary antibody mouse anti-FLAG (1:1,000, Sigma #F1804) overnight at 4 °C with gentle rocking to detect FLAG-tagged viral RNRs. The next day, cells were incubated with secondary antibody anti-mouse Alexa Fluor 594 (1:1,000, Invitrogen #A-11005) for 2 hours at room temperature protected from light and nuclei were counterstained with Hoechst 33342 (ThermoFisher #62249) for 10 minutes. Representative images were acquired using a Cytation 5 Cell Imaging Multi-Mode Reader (BioTek) at 20× magnification.

In **Figures 2-4**, approximately 5,000 HeLa cells were seeded into 96-well imaging plates (Corning #3904) and transfected the next day with 15 ng pcDNA5/TO-A3-EGFP and 30 ng pcDNA4/TO-RNR-3×FLAG. Cells were fixed approximately 24 hours after transfection. Permeabilization, antibody staining, and nuclei counterstaining with Hoechst 33342 were performed as described above. Plates were imaged using a Cytation 5 Cell Imaging Multi-Mode Reader (BioTek) at 20× magnification using the automated imaging function to capture 50 non-overlapping images per well. Background subtraction was performed using Fiji [75] and images were analyzed using CellProfiler v4.2 [76]. A pipeline was generated to identify individual nuclei, trace cell boundaries, segment the cytoplasmic/nuclear compartments, and take measurements of intensity, texture, and correlation for the Hoechst (nuclear stain), RNR, and A3 channels within each subcellular compartment. At least 100 cells were measured for each well. All measurements were exported to an SQLite database and CellProfiler Analyst v3.0 [77] was used for scoring

phenotypes by machine learning in order to quantify the number of cells showing relocalized versus non-relocalized A3 within the population of RNR-expressing cells in each well. The percentages of cells with relocalized A3 in **Figure 2D**, **Figure 3D**, and **Figure 4C** were calculated for n = 3 independent experimental replicates (i.e., data generated from the analysis of cells seeded, transfected, and imaged on three different days). The ridgeline plots in **Figure 2E**, **Figure 3E**, and **Figure 4D** were generated in RStudio with the ggplot2 package [78] using the Pearson correlation coefficient measurements for the A3 and Hoechst channels within each segment cell nucleus.

In **Figure 7B** and **Figure 8C**, approximately 5,000 HeLa cells were seeded into 96-well imaging plates (Ibidi #89626) and transfected the next day with 15 ng pcDNA5/TO-A3-EGFP and 30 ng pcDNA4/TO-RNR-3×FLAG. Cells were fixed approximately 24 hours after transfection. Permeabilization, antibody staining, and nuclei counterstaining with Hoechst 33342 were performed as described above. Plates were imaged using an Eclipse Ti2-E inverted microscope (Nikon) at 20× magnification.

ssDNA deaminase activity assays

In **Figure 5A**, **Figure 5C**, and **Figure 6E**, approximately 250,000 HeLa T-REx cells stably expressing doxycycline-inducible A3B-EGFP or A3A-EGFP (described above) were seeded into 6-well plates and transfected the next day with 200 ng pcDNA4/TO-RNR-3×FLAG using TransIT-LT1 (Mirus #2304). A3B-EGFP or A3A-EGFP expression was induced 6 hours after transfection with 50 ng/mL doxycycline. Approximately 24 hours after induction, cells were harvested, resuspended in 100 µL of reaction buffer (25 mM HEPES, 15 mM EDTA, 10% glycerol, 1 tablet of Sigma-Aldrich cOmplete Protease Inhibitor Cocktail), and subjected to freeze-thaw. Whole cell lysates were then centrifuged at 15,000 × g for 15 minutes and the clarified supernatant was

transferred to a new tube. Soluble lysates were incubated with 100 µg/mL RNAse A at room temperature for 1 hour. A 5 µl aliquot was removed from each sample for immunoblots using mouse anti-GFP (1:5,000, Clontech #632381) to detect A3-EGFP, rabbit anti-FLAG (1:5,000, Sigma-Aldrich #F7425) to detect RNR-3×FLAG, and rabbit anti-GAPDH (1:5,000, Proteintech #60004) to detect GAPDH (loading control). Secondary antibodies used were IRDye 800CW anti-rabbit (LI-COR #926-32211) and HRP-linked anti-mouse (Cell Signaling #7076), both at 1:10,000 dilution. 5 µl of soluble lysate were incubated at 37 °C for 1 hour with 0.7 µM of a fluorescent ssDNA substrate (5′-ATT ATT ATT ATT CAA ATG GAT TTA TTT ATT TAT TTA TTT ATT T-fluorescein-3′) and 2.5 U of UDG (NEB #M0280) in a total reaction volume of 10 µl (diluted in reaction buffer). NaOH was then added to the reaction mix to a final concentration of 100 mM and samples were incubated at 98 °C for 10 minutes. The reaction was then mixed with 11 µl 2× formamide buffer (80% formamide, 1×TBE, bromophenol blue, and xylene cyanol) and reaction products were separated on a 15% TBE-urea PAGE gel. Separated DNA fragments were visualized on a Typhoon FLA 7000 scanner (GE Healthcare) on fluorescence mode.

In **Figure 5D** and **Figure 8D**, approximately 250,000 293T cells were seeded in 6-well plates and transfected the next day with 100 ng pcDNA5/TO-A3-EGFP and 200 ng pcDNA4/TO-EBV-BORF2-3×FLAG. Cells were harvested approximately 24 hours after transfection and the DNA deaminase activity assay proceeded as described above.

Protein structure modeling

Protein structure models were generated with RoseTTAFold [34] using the ROBETTA server followed by simple all-atom refinement of the top-ranked models using the Relax application in the ROSIE 2 server [36]. To model the structure of the ORF61-A3Bctd complex

(**Figure 1**), RoseTTAFold-predicted models of KSHV ORF61 and A3Bctd were aligned to the cryo-EM BORF2-A3Bctd complex (pdb: 7RW6) using PyMOL [35] to arrange the starting docking pose and local protein-protein docking was performed using the RosettaDock 4.0 server [36, 37]. The 10,000 generated structures were sorted based on Critical Assessment of Predicted Interactions (CAPRI) score [79] and subsequently ranked according to their interface energy score (**Figure 1—figure supplement 1B**). All protein structure figures were made using PyMOL [35].

Herpesvirus RNR phylogenetic trees generation

Amino acid sequences of the large RNR subunits from the following herpesviruses were obtained from the NCBI Protein RefSeq database: HSV-1 (YP_009137114.1), HSV-2 (YP_009137191.1), VZV (NP_040142.1), EBV (YP_401655.1), HCMV (YP_081503.1), HHV-6A (NP_042921.1), HHV-6B (NP_050209.1), HHV-7 (YP_073768.1), KSHV (YP_001129418.1), McHV-4 (YP_067953.1), CalHV-3 (NP_733909.1), McHV-5 (NP_570809.1), McHV-11 (AAT00096.1), McHV-12 (AJE29717.1), AtHV-3 (AAC95585.1), SaHV-2 (NP_040263.1), ColHV-1 (QDQ69272.1), and SgHV-1 (UNP64460.1). To generate the phylogenetic trees in **Figure 6B** and **Figure 8E**, protein sequences were aligned using MUSCLE [80] and maximum likelihood phylogenies were inferred using PhyML 3.0 [81] using the Smart Model Selection (SMS) tool [82] and 100 bootstraps.

Ancestral APOBEC3B sequence reconstruction

To generate the phylogenetic tree in **Figure 8A**, primate A3B amino acid sequences were aligned using MUSCLE [80] (**Figure 8—figure supplement 1**) and a maximum likelihood phylogeny was inferred using PhyML 3.0 [81] using the Smart Model Selection (SMS) tool [82]

and 100 bootstraps. Reconstruction of the ancestral A3B protein sequence was performed with GRASP [44] using the sequence alignment and phylogenetic tree above.

Positive Selection Analysis

Primate *A3B* DNA sequences were aligned using ClustalOmega [83] and a maximum likelihood phylogeny was inferred using PhyML 3.0 [81] using the Smart Model Selection (SMS) tool [82] and 100 bootstraps. Codon sites under episodic diversifying selection were identified with the HyPhy 2.5 software package [84] using the mixed effects model of evolution algorithm (MEME) [46] with 100 bootstraps and p-value threshold of 0.1 (**Figure 8—figure supplement 4A-B**).

Acknowledgements

We thank members of the Harris lab for thoughtful comments. These studies were supported in part by grants to RSH from the National Institute for Allergy and Infectious Diseases (NIAID) R37-AI064046 and the National Cancer Institute (NCI) P01-CA234228. Salary support for SNM was provided by NIAID F31-AI161910 and subsequently by a grant to the University of Minnesota (SNM and RSH) from the Howard Hughes Medical Institute through the James H. Gilliam Fellowships for Advanced Study program. Salary support for JTB was provided by NIAID F32-AI147813. Salary support for NMS and AAA was provided in part by NIAID R56-AI150402 and NIH T32-AI83196, respectively. RSH is the Ewing Halsell President's Council Distinguished Chair and an Investigator of the Howard Hughes Medical Institute. The authors have no competing interests to declare.

References

1. Duggal NK, Emerman M. Evolutionary conflicts between viruses and restriction factors shape immunity. Nature Reviews Immunology. 2012;12(10):687-95. Epub 2012/09/14. doi: 10.1038/nri3295. PubMed PMID: 22976433; PubMed Central PMCID: PMCPMC3690816.
2. Harris RS, Dudley JP. APOBECs and virus restriction. Virology. 2015;479-480:131-45. Epub 2015/03/26. doi: 10.1016/j.virol.2015.03.012. PubMed PMID: 25818029; PubMed Central PMCID: PMCPMC4424171.
3. Simon V, Bloch N, Landau NR. Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. Nature Immunology. 2015;16(6):546-53.
4. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature. 2002;418(6898):646-50. Epub 2002/07/14. doi: 10.1038/nature00939. PubMed PMID: 12167863.
5. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH. DNA deamination mediates innate immunity to retroviral infection. Cell. 2003;113(6):803-9. doi: 10.1016/s0092-8674(03)00423-9. PubMed PMID: 12809610.
6. Wiegand HL, Doehle BP, Bogerd HP, Cullen BR. A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. The EMBO Journal. 2004;23(12):2451-8. doi: 10.1038/sj.emboj.7600246. PubMed Central PMCID: PMCPMC423288.
7. Sasada A, Takaori-Kondo A, Shirakawa K, Kobayashi M, Abudu A, Hishizawa M, Imada K, Tanaka Y, Uchiyama T. APOBEC3G targets human T-cell leukemia virus type 1. Retrovirology. 2005;2(1):32. doi: 10.1186/1742-4690-2-32.

8. Dang Y, Wang X, Esselman WJ, Zheng Y-H. Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. Journal of Virology. 2006;80(21):10522-33. doi: 10.1128/JVI.01123-06.

9. Esnault C, Heidmann O, Delebecque F, Dewannieux M, Ribet D, Hance AJ, Heidmann T, Schwartz O. APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature. 2005;433(7024):430-3. doi: 10.1038/nature03238.

10. Lee YN, Malim MH, Bieniasz PD. Hypermutation of an ancient human retrovirus by APOBEC3G. Journal of Virology. 2008;82(17):8762-70.

11. Turelli P, Mangeat B, Jost S, Vianin S, Trono D. Inhibition of hepatitis B virus replication by APOBEC3G. Science. 2004;303(5665):1829. doi: 10.1126/science.1092066. PubMed PMID: 15031497.

12. Suspène R, Guétard D, Henry M, Sommer P, Wain-Hobson S, Vartanian J-P. Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. Proceedings of the National Academy of Sciences. 2005;102(23):8321-6. doi: 10.1073/pnas.0408223102.

13. Vartanian J-P, Guétard D, Henry M, Wain-Hobson S. Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. Science. 2008;320(5873):230-3.

14. Peretti A, Geoghegan EM, Pastrana DV, Smola S, Feld P, Sauter M, Lohse S, Ramesh M, Lim ES, Wang D. Characterization of BK polyomaviruses from kidney transplant recipients suggests a role for APOBEC3 in driving in-host virus evolution. Cell Host & Microbe. 2018;23(5):628-35. e7.

15. Cheng AZ, Yockteng-Melgar J, Jarvis MC, Malik-Soni N, Borozan I, Carpenter MA, Mccann

JL, Ebrahimi D, Shaban NM, Marcon E, Greenblatt J, Brown WL, Frappier L, Harris RS. Epstein–Barr virus BORF2 inhibits cellular APOBEC3B to preserve viral genome integrity. Nature Microbiology. 2019;4(1):78-88. doi: 10.1038/s41564-018-0284-6.

16. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. Nature Reviews Immunology. 2004;4(11):868-77. doi: 10.1038/nri1489.

17. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature. 2003;424(6944):99-103. doi: 0.1038/nature01709.

18. Navarro F, Bollman B, Chen H, König R, Yu Q, Chiles K, Landau NR. Complementary function of the two catalytic domains of APOBEC3G. Virology. 2005;333(2):374-86. doi: 10.1016/j.virol.2005.01.011.

19. Malim MH, Bieniasz PD. HIV restriction factors and mechanisms of evasion. Cold Spring Harbor Perspectives in Medicine. 2012;2(5):a006940.

20. Hu Y, Knecht KM, Shen Q, Xiong Y. Multifaceted HIV-1 Vif interactions with human E3 ubiquitin ligase and APOBEC3s. The FEBS journal. 2021;288(11):3407-17.

21. Uriu K, Kosugi Y, Ito J, Sato K. The battle between retroviruses and APOBEC3 genes: its past and present. Viruses. 2021;13(1):124.

22. Kitamura S, Ode H, Iwatani Y. Structural features of antiviral APOBEC3 proteins are linked to their functional activities. Frontiers in Microbiology. 2011;2:258.

23. Aydin H, Taylor MW, Lee JE. Structure-guided analysis of the human APOBEC3-HIV restrictome. Structure. 2014;22(5):668-84.

24. Shapiro M, Krug LT, MacCarthy T. Mutational pressure by host APOBEC3s more strongly affects genes expressed early in the lytic phase of herpes simplex virus-1 (HSV-1) and

human polyomavirus (HPyV) infection. PLoS Pathogens. 2021;17(4):e1009560.

25. Shi K, Carpenter MA, Banerjee S, Shaban NM, Kurahashi K, Salamango DJ, McCann JL, Starrett GJ, Duffy JV, Demir Ö, Amaro RE, Harki DA, Harris RS, Aihara H. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. Nature Structural & Molecular Biology. 2017;24(2):131-9. Epub 2016/12/19. doi: 10.1038/nsmb.3344. PubMed PMID: 27991903; PubMed Central PMCID: PMCPMC5296220.

26. Kouno T, Silvas TV, Hilbert BJ, Shandilya S, Bohn MF, Kelch BA, Royer WE, Somasundaran M, Kurt Yilmaz N, Matsuo H. Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. Nature Communications. 2017;8(1):1-8.

27. Maiti A, Myint W, Kanai T, Delviks-Frankenberry K, Sierra Rodriguez C, Pathak VK, Schiffer CA, Matsuo H. Crystal structure of the catalytic domain of HIV-1 restriction factor APOBEC3G in complex with ssDNA. Nature Communications. 2018;9(1):1-11.

28. Silvas TV, Schiffer CA. APOBEC3s: DNA-editing human cytidine deaminases. Protein Science. 2019;28(9):1552-66.

29. Maiti A, Hou S, Schiffer CA, Matsuo H. Interactions of APOBEC3s with DNA and RNA. Current Opinion in Structural Biology. 2021;67:195-204.

30. Shaban NM, Yan R, Shi K, Moraes SN, Cheng AZ, Carpenter MA, McLellan JS, Yu Z, Harris RS. Cryo-EM structure of the EBV ribonucleotide reductase BORF2 and mechanism of APOBEC3B inhibition. Science advances. 2022;8(17):eabm2827.

31. Stewart JA, Holland TC, Bhagwat AS. Human Herpes Simplex Virus-1 depletes APOBEC3A from nuclei. Virology. 2019;537:104-9. doi: 10.1016/j.virol.2019.08.012.

32. Cheng AZ, Moraes SN, Attarian C, Yockteng-Melgar J, Jarvis MC, Biolatti M, Galitska G, Dell'Oste V, Frappier L, Bierle CJ, Rice SA, Harris RS. A Conserved Mechanism of APOBEC3 Relocalization by Herpesviral Ribonucleotide Reductase Large Subunits. Journal of Virology. 2019;93(23). doi: 10.1128/jvi.01539-19.

33. Cheng AZ, Moraes SN, Shaban NM, Fanunza E, Bierle CJ, Southern PJ, Bresnahan WA, Rice SA, Harris RS. APOBECs and Herpesviruses. Viruses. 2021;13(3):390. doi: 10.3390/v13030390.

34. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer RD, Millán C, Park H, Adams C, Glassman CR, DeGiovanni A, Pereira JH, Rodrigues AV, Dijk AAv, Ebrecht AC, Opperman DJ, Sagmeister T, Buhlheller C, Pavkov-Keller T, Rathinaswamy MK, Dalwadi U, Yip CK, Burke JE, Garcia KC, Grishin NV, Adams PD, Read RJ, Baker D. Accurate prediction of protein structures and interactions using a three-track neural network. Science. 2021;373(6557):871-6. doi: doi:10.1126/science.abj8754.

35. DeLano WL. The PyMOL molecular graphics system. http://www pymol org. 2002.

36. Lyskov S, Chou F-C, Conchúir SÓ, Der BS, Drew K, Kuroda D, Xu J, Weitzner BD, Renfrew PD, Sripakdeevong P. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). PloS one. 2013;8(5):e63906.

37. Marze NA, Roy Burman SS, Sheffler W, Gray JJ. Efficient flexible backbone protein–protein docking for challenging targets. Bioinformatics. 2018;34(20):3461-9.

38. Münk C, Willemsen A, Bravo IG. An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals. BMC Evolutionary Biology. 2012;12(1):71. doi: 10.1186/1471-2148-12-71.

39. Larue RS, Jonsson SR, Silverstein KA, Lajoie M, Bertrand D, El-Mabrouk N, Hotzel I, Andresdottir V, Smith TP, Harris RS. The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. BMC Molecular Biology. 2008;9(1):104. doi: 10.1186/1471-2199-9-104.

40. Ito J, Gifford RJ, Sato K. Retroviruses drive the rapid evolution of mammalianAPOBEC3genes. Proceedings of the National Academy of Sciences. 2020;117(1):610-8. doi: 10.1073/pnas.1914183116.

41. Conticello SG. The AID/APOBEC family of nucleic acid mutators. Genome Biology. 2008;9(6):229. doi: 10.1186/gb-2008-9-6-229.

42. Uriu K, Kosugi Y, Suzuki N, Ito J, Sato K. Elucidation of the Complicated Scenario of Primate APOBEC3 Gene Evolution. Journal of Virology. 2021;95(12):e00144-21.

43. Larue RS, Andrésdóttir V, Blanchard Y, Conticello SG, Derse D, Emerman M, Greene WC, Jónsson SNR, Landau NR, Löchelt M, Malik HS, Malim MH, MüNk C, O'Brien SJ, Pathak VK, Strebel K, Wain-Hobson S, Yu X-F, Yuhki N, Harris RS. Guidelines for Naming Nonprimate APOBEC3 Genes and Proteins. Journal of Virology. 2009;83(2):494-7. doi: 10.1128/jvi.01976-08.

44. Foley G, Mora A, Ross CM, Bottoms S, Sützl L, Lamprecht ML, Zaugg J, Essebier A, Balderson B, Newell R. Engineering indel and substitution variants of diverse and ancient enzymes using Graphical Representation of Ancestral Sequence Predictions (GRASP). 2019.

45. Dhingra A, Ganzenmueller T, Hage E, Suárez NM, Mätz-Rensing K, Widmer D, Pöhlmann S, Davison AJ, Schulz TF, Kaul A. Novel Virus Related to Kaposi's Sarcoma–Associated

Herpesvirus from Colobus Monkey. Emerging infectious diseases. 2019;25(8):1548.

46. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting Individual Sites Subject to Episodic Diversifying Selection. PLoS Genetics. 2012;8(7):e1002764. doi: 10.1371/journal.pgen.1002764.

47. Richards C, Albin JS, Demir Ö, Shaban NM, Luengas EM, Land AM, Anderson BD, Holten JR, Anderson JS, Harki DA. The binding interface between human APOBEC3F and HIV-1 Vif elucidated by genetic and computational approaches. Cell reports. 2015;13(9):1781-8.

48. Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. Nature Reviews Cancer. 2004;4(10):757-68. doi: Epstein–Barr virus: 40 years on.

49. Koning FA, Newman ENC, Kim E-Y, Kunstman KJ, Wolinsky SM, Malim MH. Defining APOBEC3 Expression Patterns in Human Tissues and Hematopoietic Cell Subsets. Journal of Virology. 2009;83(18):9474-85. doi: doi:10.1128/JVI.01089-09.

50. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. Nucleic Acids Research. 2010;38(13):4274-84. doi: 10.1093/nar/gkq174.

51. Blasig C, Zietz C, Haar B, Neipel F, Esser S, Brockmeyer NH, Tschachler E, Colombini S, Ensoli B, Stürzl M. Monocytes in Kaposi's sarcoma lesions are productively infected by human herpesvirus 8. Journal of Virology. 1997;71(10):7963-8. doi: 10.1128/jvi.71.10.7963-7968.1997.

52. Caselli E, Galvan M, Cassai E, Caruso A, Sighinolfi L, Di Luca D. Human herpesvirus 8 enhances human immunodeficiency virus replication in acutely infected cells and induces

reactivation in latently infected cells. Blood. 2005;106(8):2790-7. doi: 10.1182/blood-2005-04-1390.

53. Wu W, Vieira J, Fiore N, Banerjee P, Sieburg M, Rochford R, Harrington W, Feuer G. KSHV/HHV-8 infection of human hematopoietic progenitor (CD34+) cells: persistence of infection during hematopoiesis in vitro and in vivo. Blood. 2006;108(1):141-51. doi: 10.1182/blood-2005-04-1697.

54. Kim I-J, Flaño E, Woodland DL, Lund FE, Randall TD, Blackman MA. Maintenance of long term γ-herpesvirus B cell latency is dependent on CD40-mediated development of memory B cells. The Journal of Immunology. 2003;171(2):886-92. doi: 10.4049/jimmunol.171.2.886.

55. Chakraborty S, Valiya Veettil M, Chandran B. Kaposi's Sarcoma Associated Herpesvirus Entry into Target Cells. Frontiers in Microbiology. 2012;3. doi: 10.3389/fmicb.2012.00006.

56. Salamango DJ, McCann JL, Demir Ö, Brown WL, Amaro RE, Harris RS. APOBEC3B Nuclear Localization Requires Two Distinct N-Terminal Domain Surfaces. Journal of Molecular Biology. 2018;430(17):2695-708. Epub 2018/05/19. doi: 10.1016/j.jmb.2018.04.044. PubMed PMID: 29787764; PubMed Central PMCID: PMCPMC6456046.

57. Xiao X, Yang H, Arutiunian V, Fang Y, Besse G, Morimoto C, Zirkle B, Chen XS. Structural determinants of APOBEC3B non-catalytic domain for molecular assembly and catalytic regulation. Nucleic Acids Research. 2017;45(12):7494-506. doi: 10.1093/nar/gkx362. PubMed PMID: 28575276; PubMed Central PMCID: PMCPMC5499559.

58. McCann JL, Cristini A, Law EK, Lee SY, Tellier M, Carpenter MA, Beghè C, Kim JJ, Jarvis

MC, Stefanovska B. R-loop homeostasis and cancer mutagenesis promoted by the DNA cytosine deaminase APOBEC3B. bioRxiv2021.

59. Venkatesan S, Rosenthal R, Kanu N, McGranahan N, Bartek J, Quezada S, Hare J, Harris R, Swanton C. Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. Annals of Oncology. 2018;29(3):563-72.

60. Olson ME, Harris RS, Harki DA. APOBEC Enzymes as Targets for Virus and Cancer Therapy. Cell Chemical Biology. 2018;25(1):36-49. Epub 2017/11/16. doi: 10.1016/j.chembiol.2017.10.007. PubMed PMID: 29153851; PubMed Central PMCID: PMCPMC5775913.

61. Hayward JA, Tachedjian M, Cui J, Cheng AZ, Johnson A, Baker ML, Harris RS, Wang L-F, Tachedjian G. Differential evolution of antiretroviral restriction factors in pteropid bats as revealed by APOBEC3 gene complexity. Molecular Biology and Evolution. 2018;35(7):1626-37.

62. Jebb D, Huang Z, Pippel M, Hughes GM, Lavrichenko K, Devanna P, Winkler S, Jermiin LS, Skirmuntt EC, Katzourakis A. Six reference-quality genomes reveal evolution of bat adaptations. Nature. 2020;583(7817):578-84.

63. Wibbelt G, Kurth A, Yasmum N, Bannert M, Nagel S, Nitsche A, Ehlers B. Discovery of herpesviruses in bats. Journal of General Virology. 2007;88(10):2651-5.

64. Watanabe S, Maeda K, Suzuki K, Ueda N, Iha K, Taniguchi S, Shimoda H, Kato K, Yoshikawa Y, Morikawa S. Novel betaherpesvirus in bats. Emerging Infectious Diseases. 2010;16(6):986.

65. Wu Z, Ren X, Yang L, Hu Y, Yang J, He G, Zhang J, Dong J, Sun L, Du J. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces.

Journal of Virology. 2012;86(20):10999-1012.

66. Anthony SJ, Epstein JH, Murray KA, Navarrete-Macias I, Zambrana-Torrelio CM, Solovyov A, Ojeda-Flores R, Arrigo NC, Islam A, Ali Khan S. A strategy to estimate unknown viral diversity in mammals. MBio. 2013;4(5):e00598-13.

67. Sasaki M, Setiyono A, Handharyani E, Kobayashi S, Rahmadani I, Taha S, Adiani S, Subangkit M, Nakamura I, Sawa H. Isolation and characterization of a novel alphaherpesvirus in fruit bats. Journal of Virology. 2014;88(17):9819-29.

68. Sano K, Okazaki S, Taniguchi S, Masangkay JS, Puentespina R, Eres E, Cosico E, Quibod N, Kondo T, Shimoda H. Detection of a novel herpesvirus from bats in the Philippines. Virus Genes. 2015;51(1):136-9.

69. Zheng X-y, Qiu M, Chen S-w, Xiao J-p, Ma L-z, Liu S, Zhou J-h, Zhang Q-h, Li X, Chen Z. High prevalence and diversity of viruses of the subfamily Gammaherpesvirinae, family Herpesviridae, in fecal specimens from bats of different species in southern China. Archives of Virology. 2016;161(1):135-40.

70. Shabman RS, Shrivastava S, Tsibane T, Attie O, Jayaprakash A, Mire CE, Dilley KE, Puri V, Stockwell TB, Geisbert TW. Isolation and characterization of a novel gammaherpesvirus from a microbat cell line. Msphere. 2016;1(1):e00070-15.

71. Escalera-Zamudio M, Rojas-Anaya E, Kolokotronis S-O, Taboada B, Loza-Rubio E, Méndez-Ojeda ML, Arias CF, Osterrieder N, Greenwood AD. Bats, primates, and the evolutionary origins and diversification of mammalian gammaherpesviruses. MBio. 2016;7(6):e01425-16.

72. James S, Donato D, de Thoisy B, Lavergne A, Lacoste V. Novel herpesviruses in neotropical bats and their relationship with other members of the Herpesviridae family. Infection,

Genetics and Evolution. 2020;84:104367.

73. Letko M, Seifert SN, Olival KJ, Plowright RK, Munster VJ. Bat-borne virus diversity, spillover and emergence. Nature Reviews Microbiology. 2020;18(8):461-71.

74. Lackey L, Law EK, Brown WL, Harris RS. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. Cell Cycle. 2013;12(5):762-72.

75. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. Nature Methods. 2012;9(7):676-82. Epub 2012/06/28. doi: 10.1038/nmeth.2019. PubMed PMID: 22743772; PubMed Central PMCID: PMCPMC3855844.

76. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A. CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics. 2021;22(1):1-11.

77. Stirling DR, Carpenter AE, Cimini BA. CellProfiler Analyst 3.0: accessible data exploration and machine learning for image analysis. Bioinformatics. 2021;37(21):3992-4.

78. Wickham H. Data analysis. ggplot2: Springer; 2016. p. 189-201.

79. Lensink MF, Nadzirin N, Velankar S, Wodak SJ. Modeling protein-protein, protein-peptide, and protein-oligosaccharide complexes: CAPRI 7th edition. Proteins: Structure, Function, and Bioinformatics. 2020;88(8):916-38.

80. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. 2004;32(5):1792-7. doi: 10.1093/nar/gkh340.

81. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms

and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology. 2010;59(3):307-21.

82. Lefort V, Longueville J-E, Gascuel O. SMS: smart model selection in PhyML. Molecular biology and evolution. 2017;34(9):2422-4.

83. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology. 2011;7(1):539.

84. Pond SLK, Muse SV. HyPhy: hypothesis testing using phylogenies. Statistical methods in molecular evolution: Springer; 2005. p. 125-81.

85. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. Molecular Biology and Evolution. 2017;34(7):1812-9. doi: 10.1093/molbev/msx116.

86. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). Nucleic acids research. 2018;46(D1):D708-D17.

87. Greensill J, Sheldon JA, Renwick NM, Beer BE, Norley S, Goudsmit J, Schulz TF. Two distinct gamma-2 herpesviruses in African green monkeys: a second gamma-2 herpesvirus lineage among old world primates? Journal of virology. 2000;74(3):1572-7.

88. Strand K, Harper E, Thormahlen S, Thouless ME, Tsai C-C, Rose T, Bosch ML. Two distinct lineages of macaque γ herpesviruses related to the Kaposi's sarcoma associated herpesvirus. Journal of clinical virology. 2000;16(3):253-69.

89. Lacoste V, Mauclère P, Dubreuil G, Lewis J, Georges-Courbot M-C, Gessain A. A novel γ2-herpesvirus of the Rhadinovirus 2 lineage in chimpanzees. Genome Research.

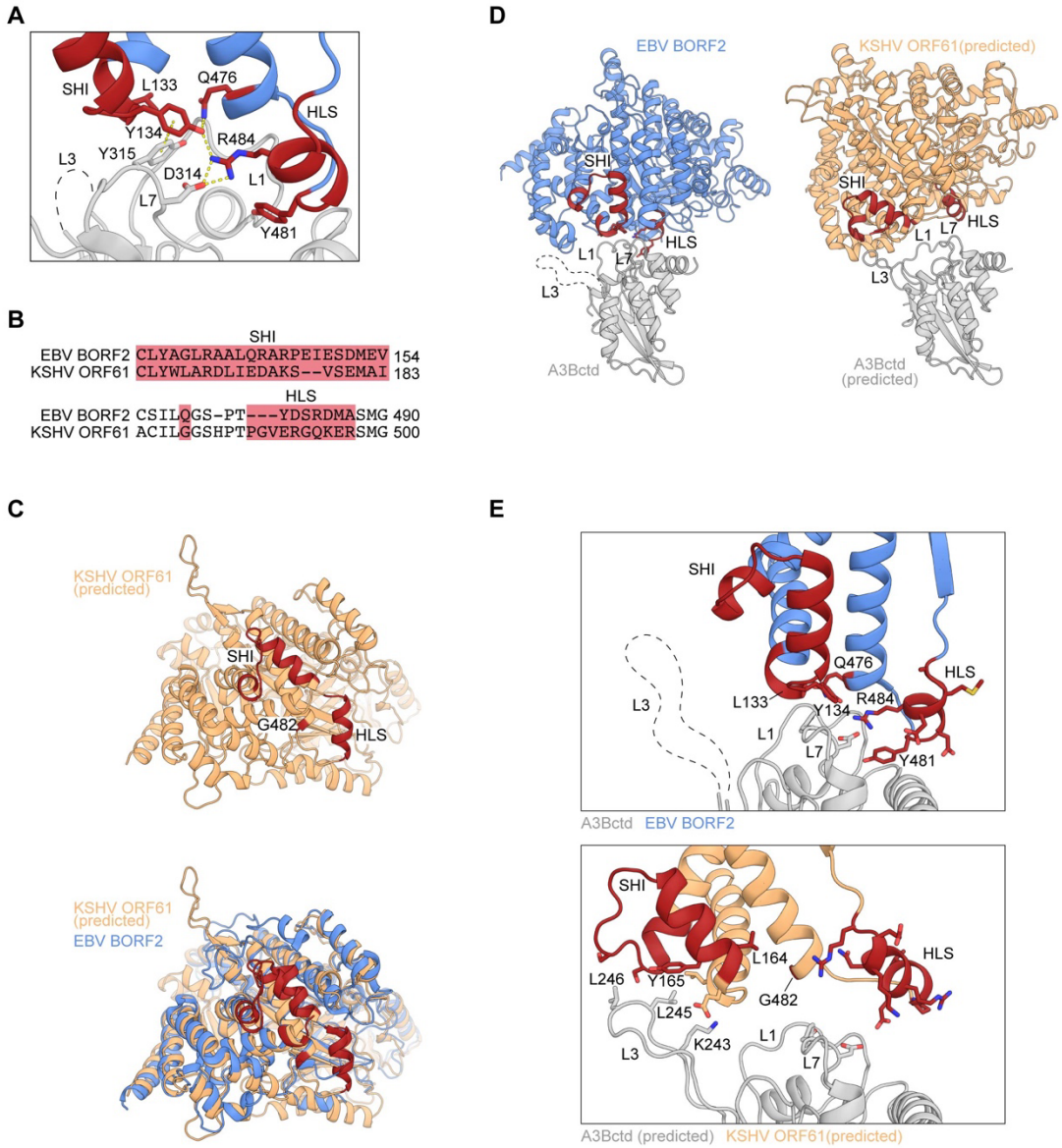
2001;11(9):1511-9.

90. Schultz ER, Rankin Jr GW, Blanc M-P, Raden BW, Tsai C-C, Rose TM. Characterization of two divergent lineages of macaque rhadinoviruses related to Kaposi's sarcoma-associated herpesvirus. Journal of virology. 2000;74(10):4919-28.

91. Whitby D, Stossel A, Gamache C, Papin J, Bosch M, Smith A, Kedes DH, White G, Kennedy R, Dittmer DP. Novel Kaposi's sarcoma-associated herpesvirus homolog in baboons. Journal of virology. 2003;77(14):8159-65.

92. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Austine-Orimoloye O, Azov AG, Barnes I, Bennett R. Ensembl 2022. Nucleic acids research. 2022;50(D1):D988-D95.

1 **Figure 1**


A
SHI
L133
Q476
HLS
Y134
R484
L3
Y315
D314
L1
L7
Y481
B
SHI
EBV BORF2 CLYAGLRAALQRARPEIESDMEV 154
KSHV ORF61 CLYWLARDLIEDAKS--VSEMAI 183
HLS
EBV BORF2 CSILQGS-PT---YDSRDMASMG 490
KSHV ORF61 ACILGGSHPTPGVERGQKERSMG 500
C
KSHV ORF61 (predicted)
SHI
G482
HLS
KSHV ORF61 (predicted)
EBV BORF2
D
EBV BORF2
KSHV ORF61(predicted)
SHI
L1
L7
HLS
L3
A3Bctd
A3Bctd (predicted)
E
SHI
Q476
HLS
L3
L133
R484
Y134
L1
L7
Y481
A3Bctd EBV BORF2
SHI
L164
HLS
L246
Y165
G482
L245
K243
L1
L3
L7
A3Bctd (predicted) KSHV ORF61(predicted)

2

Figure 1. Structural differences and similarities between EBV and KSHV RNR large subunits

(**A**) Zoom-in of the EBV BORF2-A3Bctd interface (pdb 7rw6, chains A and B) showing a network of interactions between BORF2 (blue) and A3Bctd L7 and L1 regions (gray). BORF2 short helix insertion (SHI), helical loop structure (HLS), and interacting residue Q476 are shown in red. Gray dashed line represents the L3 region not visible by cryo-EM.

(**B**) Amino acid sequence alignment of the SHI and HLS regions of EBV BORF2 and KSHV ORF61 (complete alignment in **Figure 1—figure supplement 1A**).

(**C**) RoseTTAFold-predicted structure of KSHV ORF61 (orange, above) overlayed with EBV BORF2 (pdb 7rw6, chain A; blue, below). HLS and SHI regions in both proteins and residue G482 in ORF61 are shaded red.

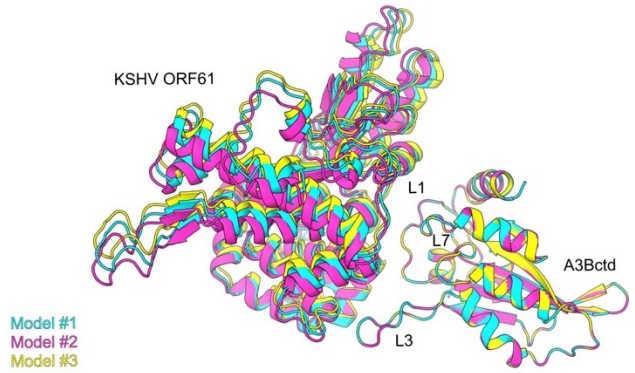
(**D-E**) Cryo-EM structure of EBV BORF2-A3Bctd (pdb 7rw6, chains A and B) and highest-ranked protein-protein docking model of KSHV ORF61-A3Bctd generated using RosettaDock (**Figure 1—figure supplement 1B**). Panel D shows the full complexes with key regions labeled. Panel E shows a zoom-in of the RNR-A3Bctd interface (BORF2, blue; ORF61, orange; A3Bctd gray; SHI and HLS regions, red). A3Bctd loops are labeled L1, L3, and L7 (gray dashed line represents the L3 region not visible by cryo-EM).

1 **Figure 1—figure supplement 1**

A

```
                                                                                                                                  SHI
EBV BORF2  MATTSHVEHEL-----------------------LSKLIDELKVKANSDPEADVLAGRLLHRLKAESVTHTVAEYLEVFSDKFYD--EEFF 66
KSHV ORF61 MSVRTFCQVHLGAVPPERRGSQVSAGQLSDFDMCAQSLIDFLKVRVGWDVRANAMAGRLWHQIMEARCPATLKQYLGIFRGVLGHRVESFI 91

EBV BORF2  QMHRDELETRVSAFAQSPAYERIVSSGYLSALRYYDTYLY--VGRSGKQESVQHFYMRLAGF----CASTTCLYAGLRAALQRARPEIESD 151
KSHV ORF61 QKNIDALEDMLCAYRRSKAYEDTLNCGYLSAVRLYDTYVLRTMGTEPVYESVAQMFMRVSVFVACQCLEHECLYWLARDLIEDAKS--VSE 180

EBV BORF2  MEVFDYYFEHLTSQTVCCSTPFMRFAGVENSTLASCILTTPDLSSEWDVTQALYRHLGRYLFQRAGVGVGVTGAGQDGKHISLLMRMINSH 242
KSHV ORF61 MAIVEYVFGYLAAQHVCCATPILRSAGVEGGQLASCFILQPSMMNEPGTLDALYHDMSPLLASKSGVGLDVTSFSHQ-KNIASCLKLVDAQ 270

EBV BORF2  VEYHNYGCKRPVSVAAYMEPWHSQIFKFLETKLPENHERCPGIFTGLFVPELFFKLFRDTPWSDWYLFDPKDAGDLERLYGEEFEREYYRL 333
KSHV ORF61 VHYFNDNNIRPVGASAYMELWHSQICDFLNAKLPENPDRCHSLFQGVCIPTLFFRMYEKDPSKLWYLFDPATAPNLIKLYGAAFDNEYERL 361

EBV BORF2  VTAGKFCGRVSIKSLMFSIVNCAVKAGSPFILLKEACNAHFWRDLQGEAMNAANLCAEVLQPSRKSVATCNLANICLPRCLVNAPLAVRAQ 424
KSHV ORF61 VRAGKYVSCMPLKSMMFTLIHTIIKTGSPYVLLKEALNEHHWTDTQGMAINCSNLCAEIVQLPGRNTSVCNLANICLPKCL----RTVESA 448
                                                                                HLS
EBV BORF2  RADTQGDELLLALPRLSVTLPGEGAVGDGFSLARLRDATQCATFVVACSILQGS-PT---YDSRDMASMGLGVQGLADVFADLGWQYTDPP 511
KSHV ORF61 RVGT------------------TDANRPFFCFEALGDAVRVAVLVINACILGGSHPTPGVERGQKERSMGIGVQGLADVFAELGYGYLDAE 521

EBV BORF2  SRSLNKEIFEHMYFTALCTSSLIGLHTRKI-FPGFKQSKYAGGWFHWHDWAGTDLS-IPREIWSRLSERIVRDGLFNSQFIALMPTSGCAQ 600
KSHV ORF61 SAELDKNIFQSMYYTAVETSHNLVLEGQGVPFHGWEVSNFAKGRFHWQTWEGEDASFVPRHRWDALGKSIAEHGIFNSQFLAVMPTAGTSQ 612

EBV BORF2  VTGCSDAFYPFYANASTKVTNKEEALRPNRSFWRHVRLDDREALNLVGGRVSCLPEALRQRYLRFQTAFDYNQEDLIQMSRDRAPFVDQSQ 691
KSHV ORF61 VTGYAESVYPFFANISSKVTNKEEVLRPNVTFFKKVLPDDLRVVRQYGGDVSTFPKHHRERYRVFLTAFDYCPFKLLDRARARAPFVDQSQ 703

EBV BORF2  SHSLFLREEDAARASTLANLLVRSYELGLKTIMYYCRIEKAADLGVMECKASAALSVPREEQNERSPAEQMPPRPMEPAQVAGPVDIMSKG 782
KSHV ORF61 SMSFFLKEDRVRNASYLRDLLLHGYRLGLKTLMYYCRVQKQSSLTALQCLADPG-SPPHSGMKQ------------DGAWLPGP-----KN 776

EBV BORF2  PGEGPGGWCVPGGLEVCYKYRQLFSEDDLLETDGFTERACESCQ 826
KSHV ORF61 PEE-----------ESC-----------------AADPECLVCQ 792
```

B



	CAPRI rank	Interface score	Total score
Model #1	2	-15.33	-1635
Model #2	2	-15.03	-1651
Model #3	2	-13.52	-1637

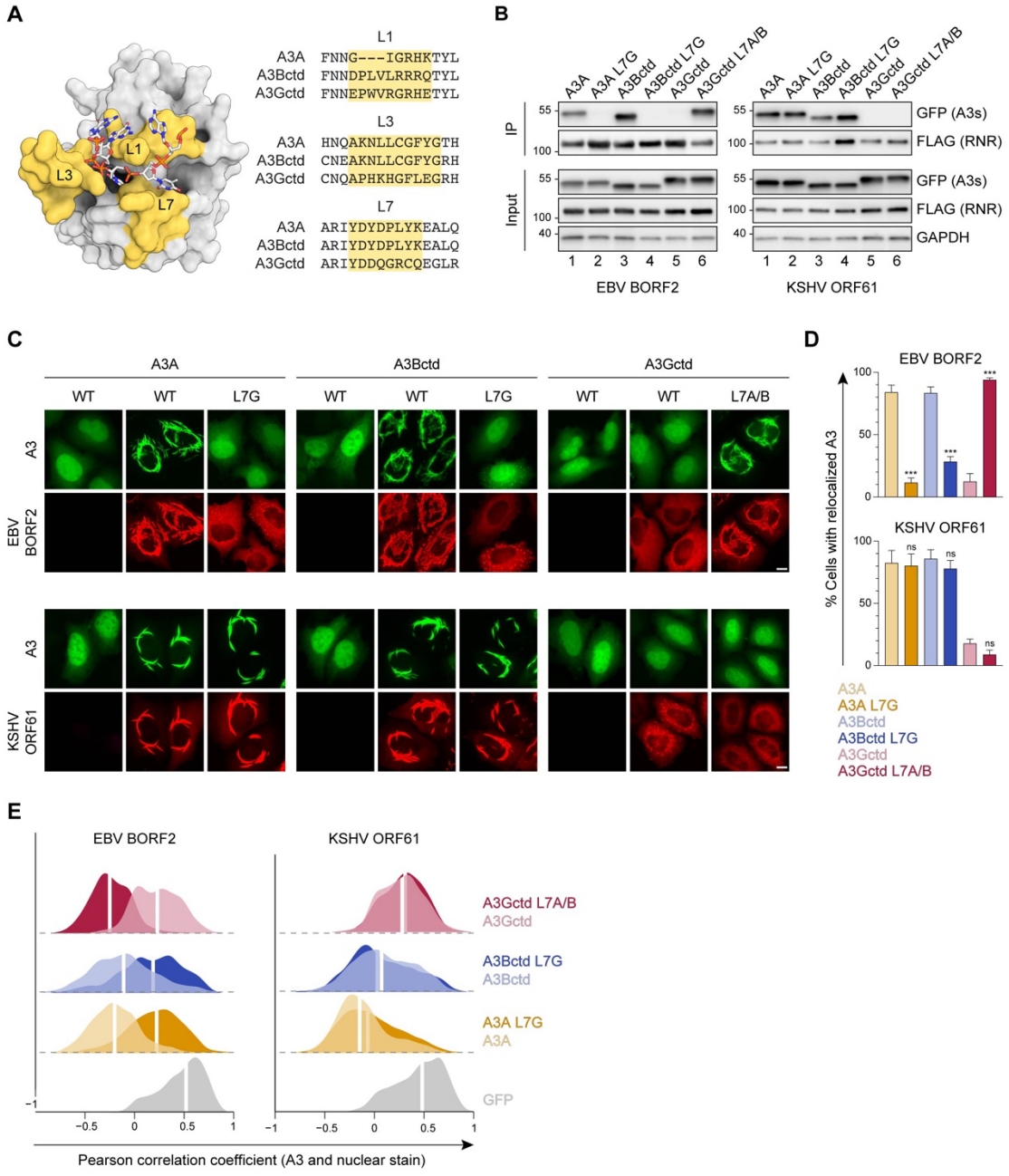
2

Figure 1—figure supplement 1

(**A**) Amino acid sequence alignment of EBV BORF2 and KSHV ORF61 with identical residues shaded yellow and the SHI and HLS regions designated by red lines.

(**B**) A superposition of the three highest-ranked protein-protein docking models of the KSHV ORF61-A3Bctd complex generated using RosettaDock (left). The CAPRI model rankings computed by RosettaDock range from 0 (incorrect model) to 3 (highest quality model; detailed in [79] with additional metrics shown to right).

1 **Figure 2**


A
L1
A3A FNNG---IGRHKTYL
A3Bctd FNNDPLVLRRRQTYL
A3Gctd FNNEPWVRGRHETYL
L3
A3A HNQAKNLLCGFYGTH
A3Bctd CNEAKNLLCGFYGRH
A3Gctd CNQAPHKHGFLEGRH
L7
A3A ARIYDYDPLYKEALQ
A3Bctd ARIYDYDPLYKEALQ
A3Gctd ARIYDDQGRCQEGLR
B
A3A
A3A L7G
A3Bctd
A3Bctd L7G
A3Gctd
A3Gctd L7A/B
IP
Input
GFP (A3s)
FLAG (RNR)
GAPDH
EBV BORF2
KSHV ORF61
C
A3A
A3Bctd
A3Gctd
WT
L7G
L7A/B
A3
EBV BORF2
KSHV ORF61
D
EBV BORF2
KSHV ORF61
% Cells with relocalized A3
ns
A3A
A3A L7G
A3Bctd
A3Bctd L7G
A3Gctd
A3Gctd L7A/B
E
EBV BORF2
KSHV ORF61
A3Gctd L7A/B
A3Gctd
A3Bctd L7G
A3Bctd
A3A L7G
A3A
GFP
Pearson correlation coefficient (A3 and nuclear stain)

2

Figure 2. Involvement of the A3B and A3A loop 7 regions in binding to gamma-herpesvirus RNRs

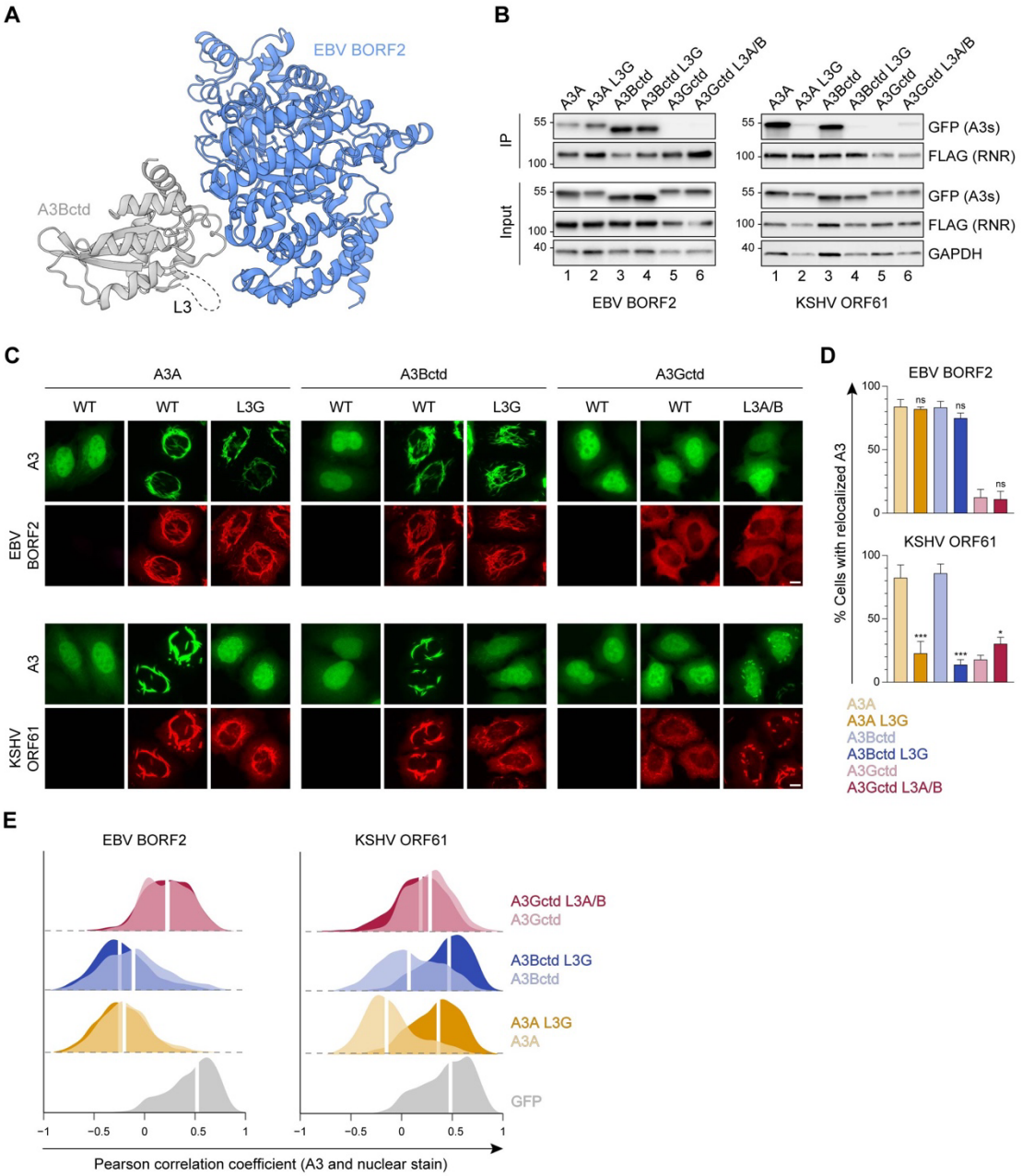
(**A**) Surface-filled representation of ssDNA bound by an A3 (pdb 5sww, chains A and E). L1, L3, and L7 regions are colored in yellow with amino acid alignments to the right.

(**B**) Co-IP of A3 L7 chimeras with EBV BORF2 and KSHV ORF61. FLAG-tagged RNR subunits were co-expressed with the indicated A3-EGFP constructs in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins.

(**C**) Representative IF microscopy images of HeLa cells co-transfected with FLAG-tagged EBV BORF2 or KSHV ORF61 (red) and the indicated A3-EGFP constructs (green). Scale = 10 μm.

(**D-E**) Quantification of A3-EGFP relocalization in HeLa cells expressing EBV BORF2 or KSHV ORF61. Panel D shows the percentage of cells exhibiting relocalized A3-EGFP (mean ± standard deviation, $n \geq 100$ cells per condition). Statistical analyses were performed using unpaired t-tests to determine significant changes in the relocalization of L7 swapped A3s (darker shade) relative to WT (lighter shade; $n = 3$ independent experimental replicates; ns, not significant $P > 0.5$; * $P \leq 0.5$; ** $P \leq 0.01$; *** $P \leq 0.001$). Panel E shows Pearson correlation coefficient (PCC) values for A3-EGFP (or EGFP alone as a control) and Hoechst (nuclear stain). Density curves show the PCC value distribution in each condition, and the white vertical lines indicate the median PCC value ($n \geq 100$ cells per condition).

1 **Figure 3**


A
EBV BORF2
A3Bctd
L3
B
A3A
A3A L3G
A3Bctd
A3Bctd L3G
A3Gctd
A3Gctd L3A/B
IP
Input
GFP (A3s)
FLAG (RNR)
GAPDH
EBV BORF2
KSHV ORF61
C
A3A
A3Bctd
A3Gctd
WT
L3G
L3A/B
A3
EBV BORF2
KSHV ORF61
D
% Cells with relocalized A3
ns
E
Pearson correlation coefficient (A3 and nuclear stain)
GFP

2

Figure 3. Role of A3B and A3A loop 3 in binding to gamma-herpesvirus RNRs

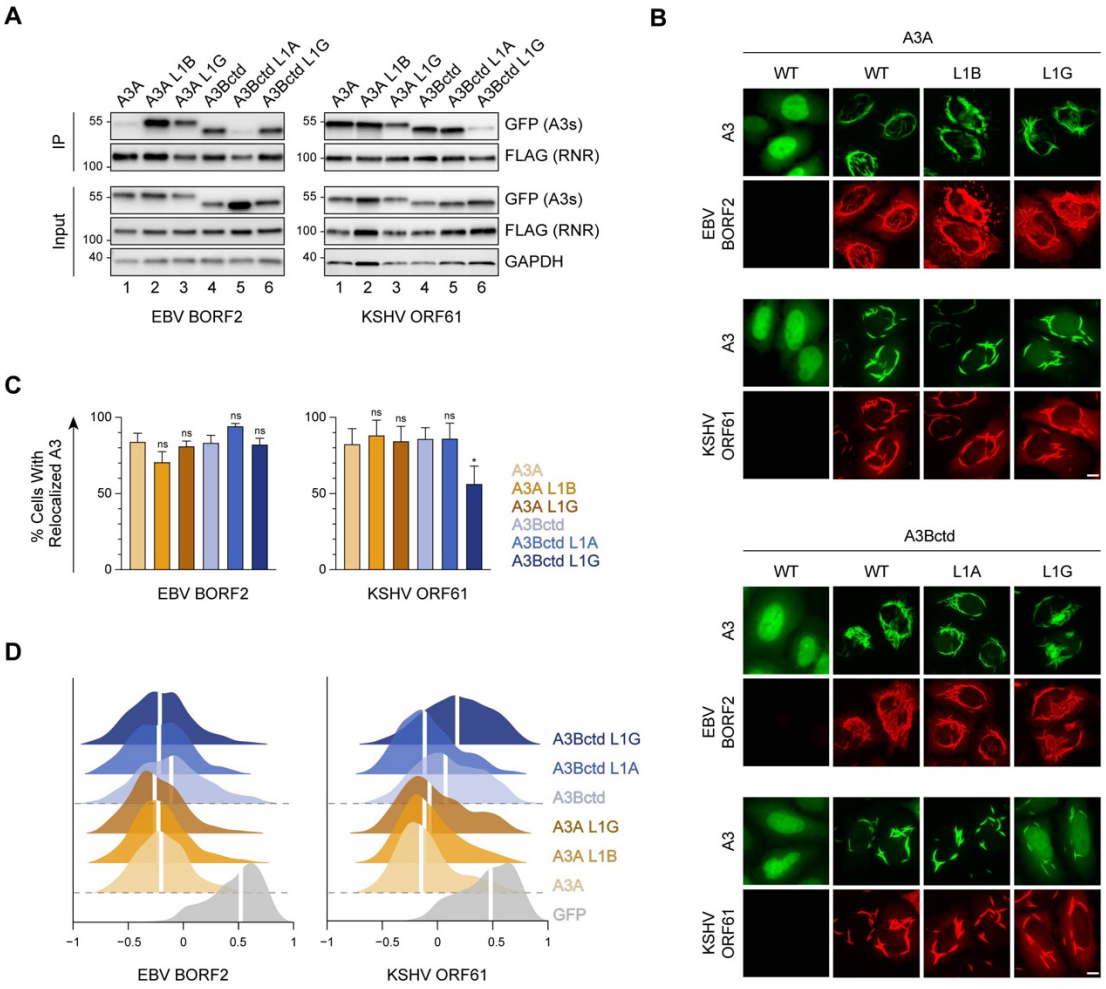
(**A**) Ribbon schematic of the EBV BORF2-A3Bctd complex (pdb 7rw6, chains A and B) with BORF2 in blue and A3Bctd in gray. Gray dashed line represents the L3 region not visible by cryo-EM.

(**B**) Co-IP of A3 L3 chimeras with EBV BORF2 and KSHV ORF61. FLAG-tagged RNR subunits were co-expressed with the indicated A3-EGFP constructs in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins.

(**C**) Representative IF microscopy images of HeLa cells co-transfected with FLAG-tagged EBV BORF2 or KSHV ORF61 (red) and the indicated A3-EGFP constructs (green). Scale = 10 μm.

(**D-E**) Quantification of A3-EGFP relocalization in HeLa cells expressing EBV BORF2 or KSHV ORF61. Panel D shows the percentage of cells exhibiting relocalized A3-EGFP (mean ± standard deviation, $n \geq 100$ cells per condition). Statistical analyses were performed using unpaired t-tests to determine significant changes in the relocalization of L3 swapped A3s (darker shade) relative to WT (lighter shade; $n = 3$ independent experimental replicates; ns, not significant $P > 0.5$; * $P \leq 0.5$; ** $P \leq 0.01$; *** $P \leq 0.001$). Panel E shows Pearson correlation coefficient (PCC) values for A3-EGFP (or EGFP alone as a control) and Hoechst (nuclear stain). Density curves show the PCC value distribution in each condition, and the white vertical lines indicate the median PCC value ($n \geq 100$ cells per condition).

1 **Figure 4**


A
A3A
A3A L1B
A3A L1G
A3Bctd
A3Bctd L1A
A3Bctd L1G
IP
Input
GFP (A3s)
FLAG (RNR)
GAPDH
EBV BORF2
KSHV ORF61
B
A3A
WT
L1B
L1G
A3
EBV BORF2
KSHV ORF61
A3Bctd
L1A
C
% Cells With Relocalized A3
ns
*
D
GFP

2

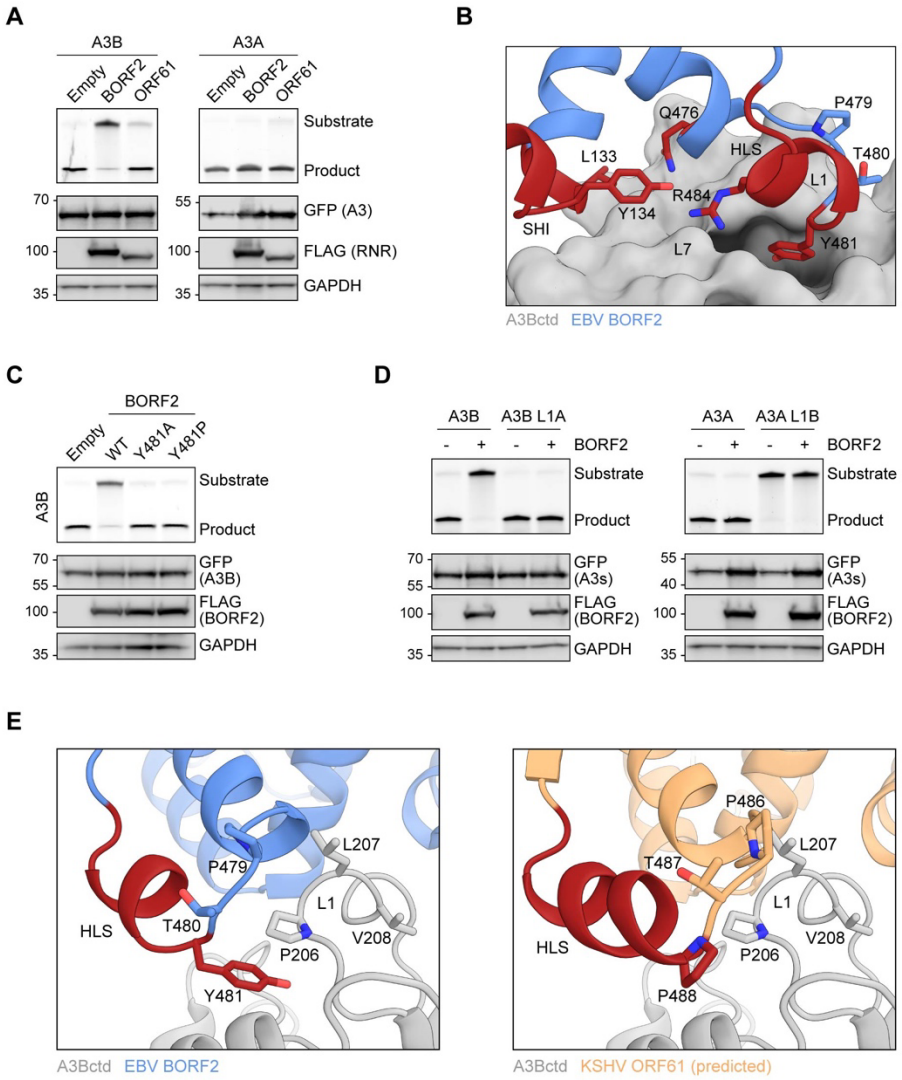
Figure 4. Role of A3B and A3A loop 1 in binding to gamma-herpesvirus RNRs

(**A**) Co-IP of A3 L1 chimeras with EBV BORF2 and KSHV ORF61. FLAG-tagged RNR subunits were co-expressed with the indicated A3-EGFP constructs in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins.

(**B**) Representative IF microscopy images of HeLa cells co-transfected with FLAG-tagged EBV BORF2 or KSHV ORF61 (red) and the indicated A3-EGFP constructs (green). Scale = 10 μm.

(**C-D**) Quantification of A3-EGFP relocalization in HeLa cells expressing EBV BORF2 or KSHV ORF61. Panel C shows the percentage of cells exhibiting relocalized A3-EGFP (mean ± standard deviation, $n \geq 100$ cells per condition). Statistical analyses were performed using unpaired t-tests to determine significant changes in the relocalization of L1 swapped A3s (darker shade) relative to WT (lighter shade; $n = 3$ independent experimental replicates; ns, not significant $P > 0.5$; * $P \leq 0.5$; ** $P \leq 0.01$; *** $P \leq 0.001$). Panel D shows Pearson correlation coefficient (PCC) values for A3-EGFP (or EGFP alone as a control) and Hoechst (nuclear stain). Density curves show the PCC value distribution in each condition, and the white vertical lines indicate the median PCC value ($n \geq 100$ cells per condition).

1 **Figure 5**


A
A3B
A3A
Empty
BORF2
ORF61
Substrate
Product
70
55
GFP (A3)
100
FLAG (RNR)
35
GAPDH
B
Q476
P479
HLS
T480
L133
L1
R484
Y134
SHI
L7
Y481
A3Bctd EBV BORF2
C
BORF2
Empty
WT
Y481A
Y481P
A3B
Substrate
Product
GFP (A3B)
FLAG (BORF2)
GAPDH
D
A3B
A3B L1A
A3A
A3A L1B
BORF2
GFP (A3s)
E
L207
P479
L1
HLS
T480
V208
P206
Y481
A3Bctd EBV BORF2
P486
T487
P488
A3Bctd KSHV ORF61 (predicted)

2

Figure 5. EBV BORF2 interacts with A3Bctd loop 1 to inhibit DNA deaminase activity.

(**A**) Upper panels show ssDNA deaminase activity of extracts from HeLa T-REx cells transfected with the indicated FLAG-tagged RNR subunits and incubated with doxycycline to induce the expression of A3B-EGFP or A3A-EGFP. Lower panels show protein expression levels in the same extracts by immunoblotting.

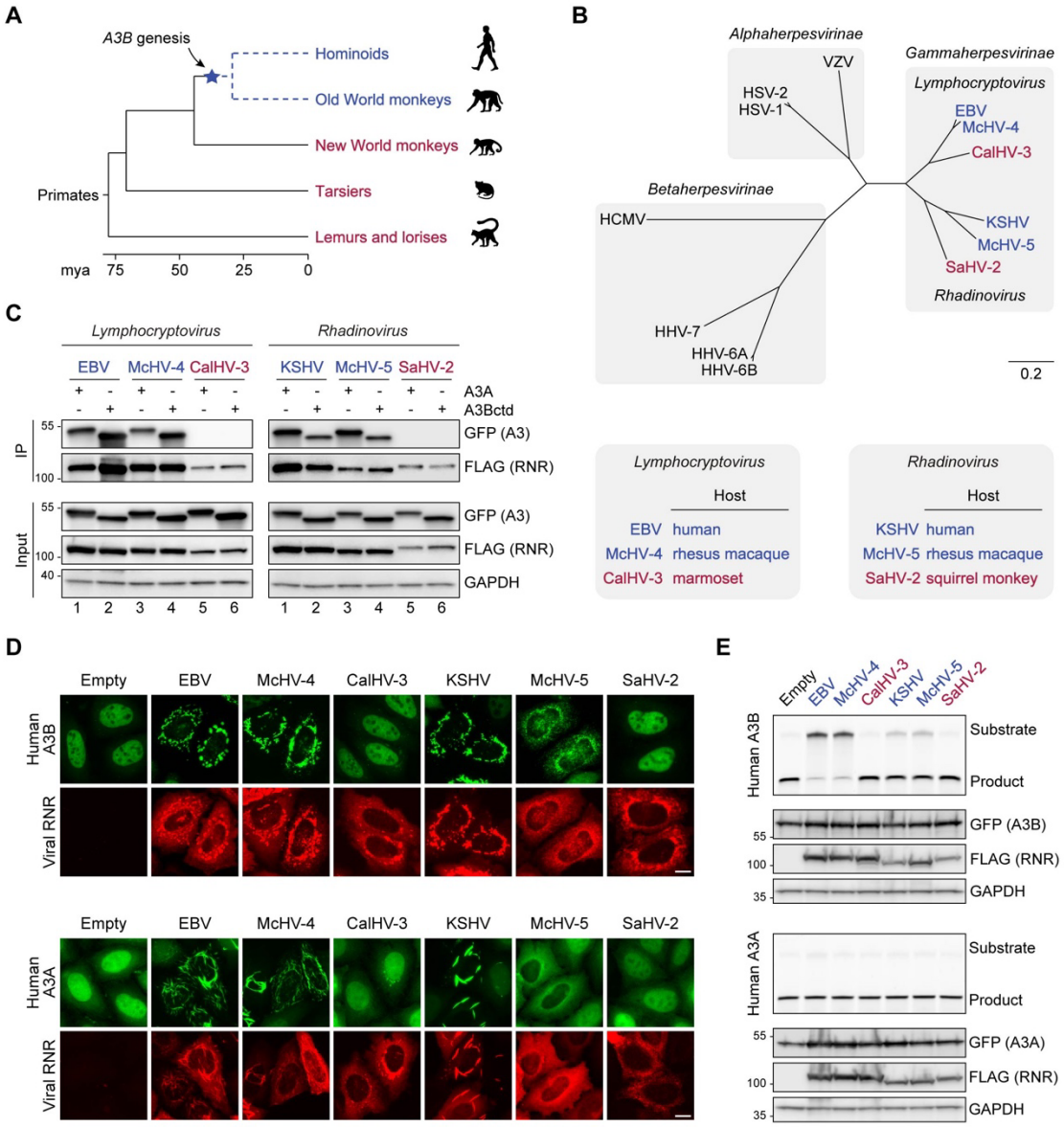
(**B**) Zoom-in of the EBV BORF2-A3Bctd interface (pdb 7rw6, chains A and B) showing a network of interactions between BORF2 (blue, ribbon representation) and A3Bctd L7 and L1 regions (gray, surface-filled representation). BORF2 SHI, HLS, and Q476 are shown in red.

(**C**) Upper panels show ssDNA deaminase activity of extracts from HeLa T-REx cells transfected the indicated FLAG-tagged EBV BORF2 constructs and incubated with doxycycline to induce the expression of A3B-EGFP. Lower panels show protein expression levels in the same extracts by immunoblotting.

(**D**) Upper panels show ssDNA deaminase activity of extracts from HeLa cells co-transfected with FLAG-tagged EBV BORF2 and the indicated A3-EGFP constructs. Lower panels show protein expression levels in the same extracts by immunoblotting.

(**E**) Ribbon schematic of the EBV BORF2-A3Bctd interface (pdb 7rw6, chains A and B) highlighting BORF2 tyrosine 481 and A3Bctd L1. BORF2 is shown in blue, the HLS region in red, and A3Bctd in gray. The right schematic shows the same view with the predicted structure of KSHV ORF61 (orange) overlayed and BORF2 residues hidden to allow better visualization (full model in **Figure 1C**).

1 **Figure 6**


A
A3B genesis
Hominoids
Old World monkeys
New World monkeys
Tarsiers
Lemurs and lorises
Primates
mya 75 50 25 0
B
Alphaherpesvirinae
VZV
HSV-2
HSV-1
Gammaherpesvirinae
Lymphocryptovirus
EBV
McHV-4
CalHV-3
Betaherpesvirinae
HCMV
KSHV
McHV-5
SaHV-2
Rhadinovirus
HHV-7
HHV-6A
HHV-6B
0.2
Lymphocryptovirus
Host
EBV human
McHV-4 rhesus macaque
CalHV-3 marmoset
Rhadinovirus
Host
KSHV human
McHV-5 rhesus macaque
SaHV-2 squirrel monkey
C
Lymphocryptovirus
Rhadinovirus
EBV McHV-4 CalHV-3
KSHV McHV-5 SaHV-2
A3A
A3Bctd
IP
Input
GFP (A3)
FLAG (RNR)
GAPDH
D
Empty EBV McHV-4 CalHV-3 KSHV McHV-5 SaHV-2
Human A3B
Viral RNR
Human A3A
E
Human A3B
Substrate
Product
GFP (A3B)
FLAG (RNR)
GAPDH
Human A3A
GFP (A3A)

2

Figure 6. The genesis of A3B may have shaped gamma-herpesvirus RNRs evolution.

(**A**) Chronogram of primates depicting the generation of *A3B* in the common ancestor of hominoids and Old World monkeys. Primate groups shown in blue are composed of species that encode *A3B* and those shown in red lack *A3B*. Primate phylogeny and divergence intervals were derived from timetree.org [85]. Mya = million years ago.

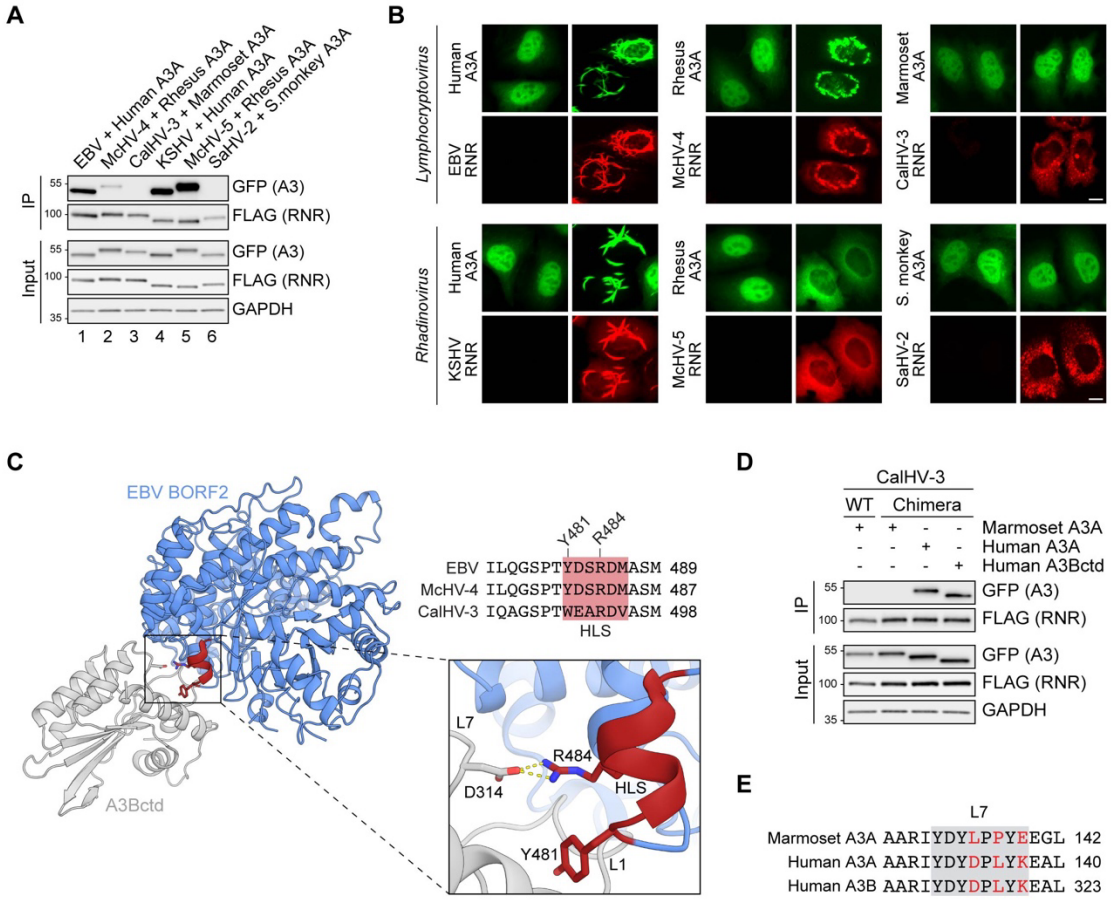
(**B**) Top: Phylogenetic tree of human herpesvirus and indicated non-human primate herpesvirus RNR large subunits. Shaded boxes indicate alpha-, beta-, and gamma-herpesvirus subfamilies. Genus classification is shown for gamma-herpesviruses. Gamma-herpesviruses that infect host species that encode *A3B* are shown in blue and those that infect host species that lack *A3B* are shown in red. Bottom: Host species infected by the indicated gamma-herpesviruses. Virus-host pairs are colored according to *A3B* status, as described for panel A.

(**C**) Co-IP of human A3A and A3Bctd with the RNR large subunits from the indicated viruses. FLAG-tagged RNR subunits were co-expressed with A3A-EGFP or A3Bctd-EGFP in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins.

(**D**) Representative IF microscopy images of HeLa T-REx cells transfected with the indicated FLAG-tagged RNR subunits (red) and incubated with doxycycline to induce the expression of human A3B-EGFP or A3A-EGFP (green). Scale = 10 μm.

(**E**) Upper panels show ssDNA deaminase activity of extracts from HeLa T-REx cells transfected with the indicated FLAG-tagged RNR subunits and incubated with doxycycline to induce the expression of A3B-EGFP or A3A-EGFP. Lower panels show protein expression levels in the same extracts by immunoblotting.

1 **Figure 7**


A
EBV + Human A3A
McHV-4 + Rhesus A3A
CalHV-3 + Marmoset A3A
KSHV + Human A3A
McHV-5 + Rhesus A3A
SaHV-2 + S.monkey A3A
IP
Input
GFP (A3)
FLAG (RNR)
GAPDH
1 2 3 4 5 6
B
Lymphocryptovirus
Rhadinovirus
Human A3A
Rhesus A3A
Marmoset A3A
S. monkey A3A
EBV RNR
McHV-4 RNR
CalHV-3 RNR
KSHV RNR
McHV-5 RNR
SaHV-2 RNR
C
EBV BORF2
A3Bctd
Y481
R484
EBV ILQGSPTYDSRDMASM 489
McHV-4 ILQGSPTYDSRDMASM 487
CalHV-3 IQAGSPTWEARDVASM 498
HLS
L7
D314
HLS
L1
D
CalHV-3
WT
Chimera
Marmoset A3A
Human A3A
Human A3Bctd
E
L7
Marmoset A3A AARIYDYLPPYEEGL 142
Human A3A AARIYDYDPLYKEAL 140
Human A3B AARIYDYDPLYKEAL 323

2

Figure 7. A short helical loop structure from EBV BORF2 enables the marmoset CalHV-3 RNR to bind to human A3B and A3A

(**A**) Co-IP of human, rhesus macaque, marmoset, and squirrel monkey A3As with the indicated gamma-herpesvirus RNR subunits. FLAG-tagged RNR subunits were co-expressed with the indicated A3-EGFP constructs in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins.

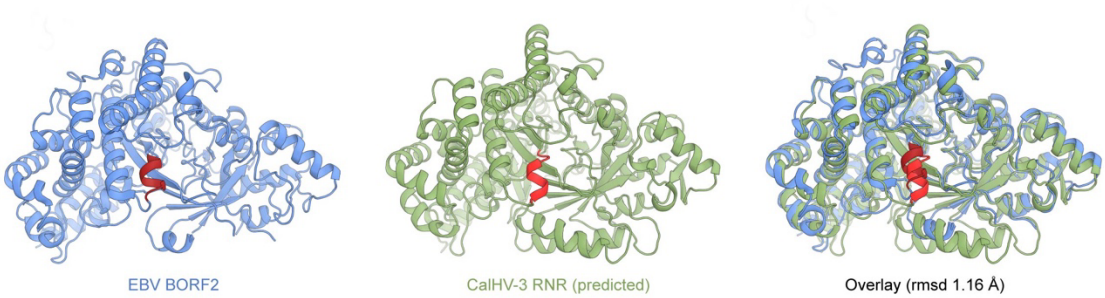
(**B**) Representative IF microscopy images of HeLa cells co-transfected with the indicated FLAG-tagged viral RNRs (red) and A3A-EGFP constructs (green). Scale = 10 μm.

(**C**) Ribbon schematic of the EBV BORF2-A3Bctd complex (pdb 7rw6, chains A and B) with BORF2 in blue and A3Bctd in gray. The HLS of EBV BORF2 is depicted in red. Bottom-right: zoom-in of the interacting surfaces highlighting BORF2 Y481 and R484 interactions with A3B loops 1 and 7, respectively. Top-right: amino acid alignment of the HLS regions of the EBV, McHV-4, and CalHV-3 RNRs (red boxes) and adjacent residues.

(**D**) Co-IP of marmoset A3A, human A3A, and human A3Bctd with WT or chimeric CalHV-3 RNR containing the HLS region of EBV BORF2. FLAG-tagged RNR subunits were co-expressed with the indicated A3-EGFP constructs in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins.

(**E**) Amino acid alignment of the loop 7 regions of marmoset A3A, human A3A, and human A3B (gray boxes) with non-identical loop 7 residues highlighted in red.

1 **Figure 7—figure supplement 1**



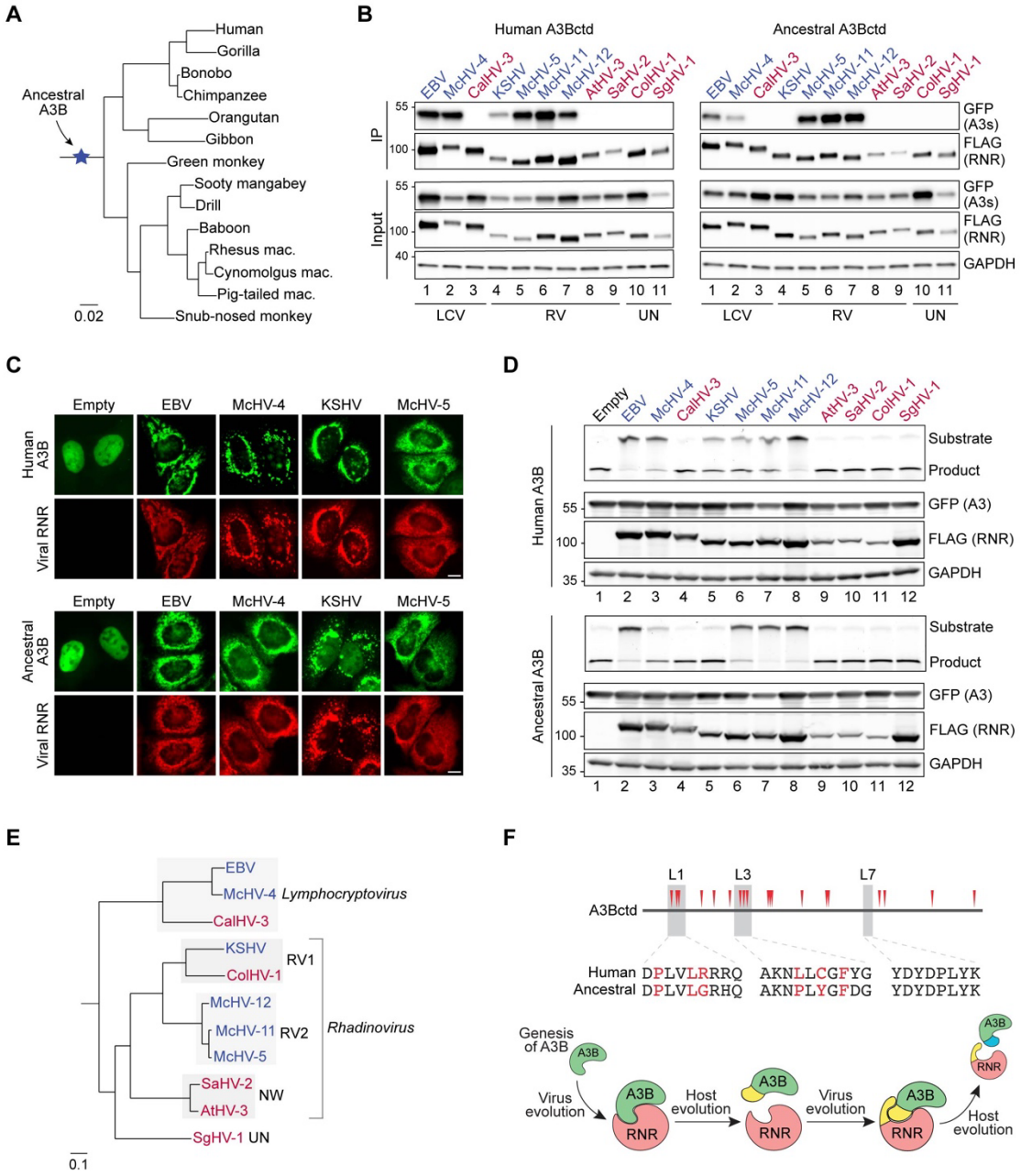
2

1 **Figure 7—figure supplement 1**

2 Ribbon schematics of the EBV BORF2 cryo-EM structure (pdb 7rw6, chain A; blue) and the
3 RoseTTAFold-predicted CalHV-3 RNR structure (green). The HLS region of each protein is
4 shaded red and a structural overlay is shown for comparison.

5

1 **Figure 8**


A
Human
Gorilla
Bonobo
Chimpanzee
Orangutan
Gibbon
Green monkey
Sooty mangabey
Drill
Baboon
Rhesus mac.
Cynomolgus mac.
Pig-tailed mac.
Snub-nosed monkey
Ancestral A3B
0.02
B
Human A3Bctd
Ancestral A3Bctd
EBV McHV-4 CalHV-3 KSHV McHV-5 McHV-11 McHV-12 AtHV-3 SaHV-2 ColHV-1 SgHV-1
IP
Input
GFP (A3s)
FLAG (RNR)
GAPDH
LCV
RV
UN
C
Empty
EBV
McHV-4
KSHV
McHV-5
Human A3B
Ancestral A3B
Viral RNR
D
Substrate
Product
GFP (A3)
FLAG (RNR)
GAPDH
E
EBV
McHV-4
CalHV-3
Lymphocryptovirus
KSHV
ColHV-1
RV1
McHV-12
McHV-11
McHV-5
RV2
Rhadinovirus
SaHV-2
AtHV-3
NW
SgHV-1 UN
0.1
F
L1
L3
L7
A3Bctd
Human DPLVLRRRQ AKNLLCGFYG YDYDPLYK
Ancestral DPLVLGRHQ AKNPLYGFDG YDYDPLYK
Genesis of A3B
Virus evolution
Host evolution
A3B
RNR

2

Figure 8. Ancestral A3B is active and antagonized by the RNRs from gamma-herpesviruses that infect *A3B*-encoding primates

(**A**) Phylogenetic tree of primate A3B sequences used in this study to reconstruct ancestral A3B (amino acid sequence alignment in **Figure 8—figure supplement 1**).

(**B**) Co-IP of human and ancestral A3Bctd with the indicated gamma-herpesvirus RNR subunits. FLAG-tagged RNR subunits were co-expressed with the indicated A3-EGFP constructs in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins. RNRs from viruses that infect host species that encode *A3B* are shown in blue, and RNRs from viruses that infect host species that lack *A3B* are shown in red. LCV = *Lymphocryptovirus*; RV = *Rhadinovirus*; UN = unclassified gamma-herpesvirus (no official genus classification from the International Committee on Taxonomy of Viruses [86]).

(**C**) Representative IF microscopy images of HeLa cells co-transfected human or ancestral A3B-EGFP (green) and the indicated FLAG-tagged viral RNRs (red). Scale = 10 μm. See **Figure 8—figure supplement 2A** for additional RNR data.

(**D**) Upper panels show ssDNA deaminase activity of extracts from HeLa cells co-transfected with human or ancestral A3B-EGFP and the indicated FLAG-tagged viral RNRs. Lower panels show protein expression levels in the same extracts by immunoblotting.

(**E**) Phylogenetic tree of primate gamma-herpesviruses RNR large subunits. *Rhadinovirus* RV1, RV2, and NW refer to primate *Rhadinovirus* lineages [87-91] (UN, unclassified).

(**F**) Top: positively selected sites in A3Bctd identified using MEME [46] with $p < 0.1$ (red triangles in schematic and red residues in loop region alignments; see **Figure 8—figure supplement 4** for details). Bottom: schematic of the proposed genetic conflict between primate gamma-herpesvirus RNRs and cellular A3B.

1 **Figure 8—figure supplement 1**

```
                                                1
Human (NM_004900.5)                             MNPQIRNPME RMYRDTFYDN FENEPILYGR SYTWLCYEVK IKRGRSNLLW DTGVFRGQVY FKPQYHAEMC FLSWFCGNQL
Gorilla (XM_031005506.1)                        .......... ....G..... .......... .......... .......... .......M. SQ.EH..... ..........
Bonobo (NM_001301766.1)                         .......... W..QR...Y. .......... .......... .R..H..... .......M. SQ.EH....Y ..........
Chimpanzee (NM_001289639.1)                     .......... W..QR...Y. .......... .......... .R..H..... .......M. SQ.EH..... ..........
Orangutan (XM_009234378.2)                      .......... ...QS...YD ........R. N......... .RKDP.K.A. .......... SQ.EH..... ..........
Gibbon (NM_001345972.1)                         .......... ....P...Y. .......CR. .......... .RKDP.K.P. .......M. .Q.E...... ..........
Green monkey (NM_001345954.1)                   .......... ...QR...YH .......... .......... .RKDP.K.P. .......M. S..EH..... ..........
Sooty mangabey (XM_012032024.1)                 .......... ...QR...YH ...K...... .......... .RKDP.K.P. Y........ S..EH..... ...R......
Drill (XM_011978846.1)                          .......... ...QR...YH ...K...... .......... .RKDP.K.P. Y........ S..EH..... ...R......
Baboon (NM_001346001.1)                         .......... ...QR...YH ...K...... .......... .RKDP.K.P. Y........ S..EH..... ...R......
Rhesus mac. (NM_001246230.2)                    .......... ...QR...YH ...K...... .......... .RKDP.K.P. .......... S..EH..... ...R......
Cynomolgus mac. (XM_005567246.2)                .......... ...QH...YH ...K...... .......... .RKDP.K.P. .......... S..EH..... ...R......
Pig-tailed mac. (KX583652.1)                    .......... ...QR...YH ...K...... .......... .RKDP.K.P. .......... S..EH..... ...R......
Snub-nosed monkey (XM_030915811.1)              .......... ...QHI..Y. .......... .......... .RKDP.K.P. .......... SE.EH....Y ..........
Ancestral                                       .......... ...QR...Y. .......... .......... .RKDP.K.P. .......M. SQ.EH..... ..........

                                                81
Human (NM_004900.5)                             PAYKCFQITW FVSWTPCPDC VAKLAEFLSE HPNVTLTISA ARLYYYWERD YRRALCRLSQ AGARVTIMDY EEFAYCWENF
Gorilla (XM_031005506.1)                        .......... ......L.. ........A. Y........T .......... .......... ....MK.... ..........
Bonobo (NM_001301766.1)                         S......... .......... ........A. .......... .......... .......... .....K.... ..........
Chimpanzee (NM_001289639.1)                     S......... .......... .....K..A. .......... .......... .......... .....K...D ..........
Orangutan (XM_009234378.2)                      S..ER..... .......... ........A. .......V.. .........G .....R..R. ...H.K.... ..........
Gibbon (NM_001345972.1)                         ....R..... .......... ...V.V..A. .......... .......K. WQ........ .....K.... ...E......
Green monkey (NM_001345954.1)                   ..H.R..... .......... ...V....A. Y......... .......T. ........R. .....K.... .........C
Sooty mangabey (XM_012032024.1)                 ....R..... ....N..... ...VI...A. .......... .......S.. WQ......R. ...G.K.... ..........
Drill (XM_011978846.1)                          ....R..... ....N..... ...VI...A. ........FT .......G.. WQ......R. .....K.... ..........
Baboon (NM_001346001.1)                         ....R..... ....N..... ...VI...A. .........T .......G.. WQ......R. .....K.... ..........
Rhesus mac. (NM_001246230.2)                    ....R..... ....N..... ...VI...A. .........T .......G.. WQ......R. .....K.... ..........
Cynomolgus mac. (XM_005567246.2)                ....R..... ....N..... ...VI...A. .........T .......G.. WQ......R. .....K.... ..........
Pig-tailed mac. (KX583652.1)                    ....R..... ....N..... .V.VI...A. .........T .......G.. WQ......R. .....K.... ..........
Snub-nosed monkey (XM_030915811.1)              ....R..... .......... ...V....T. .......... ......RG.. W.......R. .....K.... .........C
Ancestral                                       ....R..... .......... ...V....A. .......... .......... ........R. .....K.... ..........

                                                                                                        L1
                                                161
Human (NM_004900.5)                             VYNEGQQFMP WYKFDENYA- ---------F LHRTLKEILR YLMDPDTFTF NFNNDPLVLR RRQTYLCYEV ERLDNGTWVL
Gorilla (XM_031005506.1)                        .......... ......... .......... ..H....... H...E..... .......... .H........ ..........
Bonobo (NM_001301766.1)                         ......P... .....D.... .......... ........I. H......... .......... .H........ ..........
Chimpanzee (NM_001289639.1)                     ......P... .....D.... .......... ........I. H......... .......... .H........ ..........
Orangutan (XM_009234378.2)                      ......P... .....D.... .......... ..H....... Q......... ......F... .H........ .H.......K
Gibbon (NM_001345972.1)                         .....EP... .....D...F PHRFDANYA. ..H....... H......... .........FG .H........ .........K
Green monkey (NM_001345954.1)                   ...G..P... .....D.... .......... ..HK...... H........I .....LS..G .......... .........P
Sooty mangabey (XM_012032024.1)                 ....D.S... .....D.... .......... ...M...... H......... .....LS..G .H........ .HR......P
Drill (XM_011978846.1)                          ......S... .....D.... .......... ...M...... H......... .....LS..G .H........ .........P
Baboon (NM_001346001.1)                         ....D.S... .....D.... .......... ...M...... H........S .....LS.RG .H........ .H.......P
Rhesus mac. (NM_001246230.2)                    ....D.S... ....ND.... .......... ...M...... H........S .....VS.RG .H........ .........P
Cynomolgus mac. (XM_005567246.2)                ....D.S... ....ND.... .......... ...M...... H........S .....LS..G .H........ .........P
Pig-tailed mac. (KX583652.1)                    ....D.S... .....D.... .......... ...M...... H........S .L...LS.RG .H........ .........P
Snub-nosed monkey (XM_030915811.1)              A.....P... .....D.... .......... ...K...... H.....M... .....LSI.G .H........ .H.......P
Ancestral                                       ......P... .....D.... .......... ..H....... H......... ........G .H........ .........P

                                                241                L3
Human (NM_004900.5)                             MDQHMGFLCN EAKNLLCGFY GRHAELRFLD LVPSLQLDPA QIYRVTWFIS WSPCFSWGCA GEVRAFLQEN THVRLRIFAA
Gorilla (XM_031005506.1)                        ..R....... .......... .......... .......... .......... .......... .Q........ .....H....
Bonobo (NM_001301766.1)                         .......... .......... .......... .......... .......... .......... .Q........ ..........
Chimpanzee (NM_001289639.1)                     .......... .......... .......... .......... .......... .......... .Q........ ..........
Orangutan (XM_009234378.2)                      ....R.S.H. Q.R.P.Y.LD .........G .L.YW..... .......... .......... RQ........ ..........
Gibbon (NM_001345972.1)                         ....R...H. Q..DP.Y..D ...T..C..G .I.YW...L. .T........ .......... EQ........ ..M.......
Green monkey (NM_001345954.1)                   ....W..... Q.R.P.H.VD SC.V..C..S Q.S.W..... .T........ ......G... EQ........ .....C....
Sooty mangabey (XM_012032024.1)                 ....W..... Q.R.P.H.VD SC....C... Q.S.W..... .T........ .......... EQ.H...... ..........
Drill (XM_011978846.1)                          ....W..... Q.R.P.H.VD SC....C... Q.S.W..... .T........ .......... EQ.H...... ..........
Baboon (NM_001346001.1)                         ....W..... Q.....HDD. .C....C..G W...W..... .M........ ......G... EQ........ ..........
Rhesus mac. (NM_001246230.2)                    ....W..... Q...V.R.D. .C....C..G W...W..... .T........ .......... EQ........ .....H....
Cynomolgus mac. (XM_005567246.2)                ....W..... Q...V.R.D. .C....C..G W...W..... .T........ ....V..... EQ........ .....H....
Pig-tailed mac. (KX583652.1)                    ....W..... Q...VPR.D. .C.V..C..G K...W..... .T........ .......... EQ........ K.........
Snub-nosed monkey (XM_030915811.1)              ....W..... Q...P.H.VD .C....C... Q...W..... .MH....... .......... EQ........ ..........
Ancestral                                       ....R..... Q...P.Y..D .....C..G ....W..... .T........ .......... EQ........ ..........

                                                321     L7
Human (NM_004900.5)                             RIYDYDPLYK EALQMLRDAG AQVSIMTYDE FEYCWDTFVY RQGCPFQPWD GLEEHSQALS GRLRAILQNQ GN
Gorilla (XM_031005506.1)                        .......... .......... .......... .KH....... C......... .......... ...Q...... ..
Bonobo (NM_001301766.1)                         .......... .......... .......... .......... .......... .......... .......... ..
Chimpanzee (NM_001289639.1)                     .......... .......... .......... .......... .......... .......... .......... ..
Orangutan (XM_009234378.2)                      .......... .......... .......... .....N...D H......... .......... .K.Q...L.. ..
Gibbon (NM_001345972.1)                         .......... .......G.. .......... ..H......D H..R...... .......... ...Q...L.. ..
Green monkey (NM_001345954.1)                   ...N.....Q ...R...... ........E. .........D .......... ..D....... .......... ..
Sooty mangabey (XM_012032024.1)                 ........Q ...RT.Q... ........E. .........D .......... ..D....... ...Q..--. ..
Drill (XM_011978846.1)                          ........Q ...RT..... ...F....K. .........D ...R...... ..D....... ...Q...... ..
Baboon (NM_001346001.1)                         ........Q ...R...... ........E. .........D .......... ..D....... .......... ..
Rhesus mac. (NM_001246230.2)                    ........Q ...RT..... .......... .........D ...R...... ..D....... .......... ..
Cynomolgus mac. (XM_005567246.2)                ........Q ...RT..... .......... .........D ...R...... ..D....... .......... ..
Pig-tailed mac. (KX583652.1)                    ........Q ...RT..... .E......E. .........D ...R...... ..D....... .......... ..
Snub-nosed monkey (XM_030915811.1)              ........Q ...RT..... ...F....E. .........D .......... ..D....... .......... ..
Ancestral                                       .......... .......... .......... .........D .......... .......... .......... ..
```

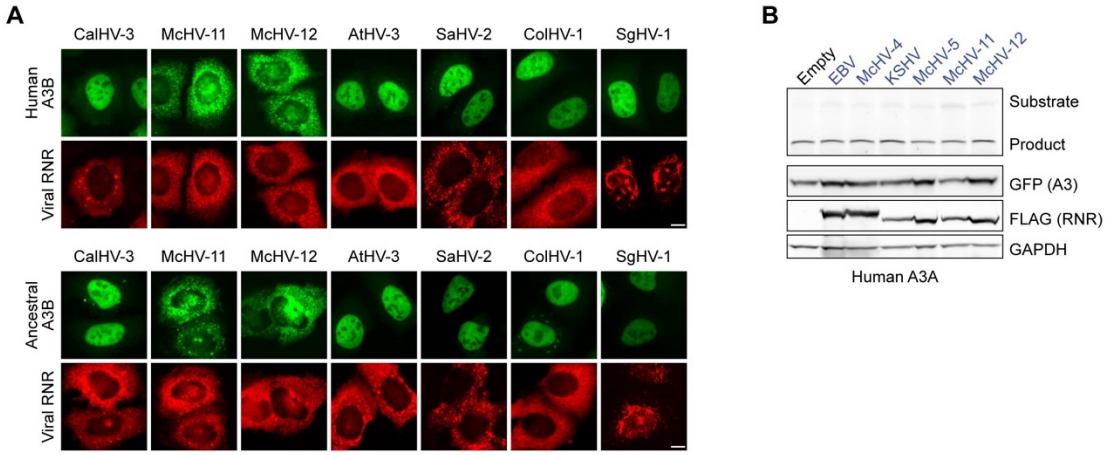
2

1 **Figure 8—figure supplement 1**

2 Amino acid sequence alignment of the indicated primate A3Bs. A3Bctd regions are shaded yellow
3 and the L1, L3, and L7 regions are highlighted with blue boxes. The predicted ancestral A3B
4 protein sequence is positioned at the bottom of the group.

5

1 **Figure 8—figure supplement 2**



2

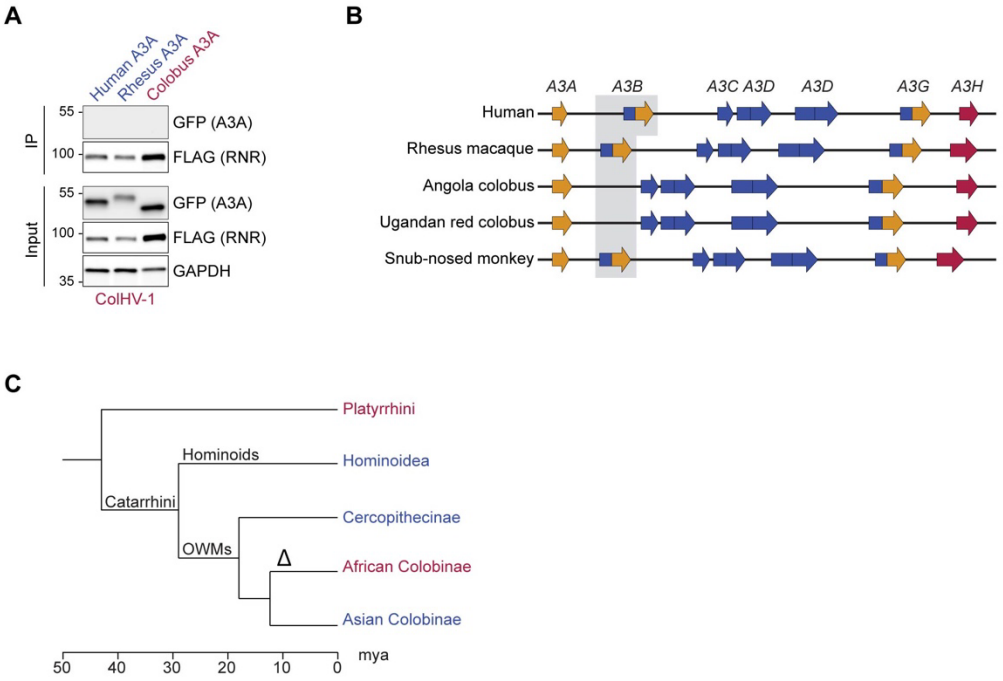
1 **Figure 8—figure supplement 2**

2 (**A**) Representative IF microscopy images of HeLa cells co-transfected human or ancestral A3B-
3 EGFP (green) and the indicated FLAG-tagged viral RNRs (red). Scale = 10 μm.

4 (**B**) Upper panels show ssDNA deaminase activity of extracts from HeLa cells co-transfected with
5 human A3A-EGFP and the indicated FLAG-tagged viral RNRs. Lower panels show protein
6 expression levels in the same extracts by immunoblotting.

7

1 **Figure 8—figure supplement 3**



2

Figure 8—figure supplement 3

(**A**) Co-IP of human, rhesus macaque, and Angola colobus A3A enzymes with the ColHV-1 RNR. FLAG-tagged ColHV-1 RNR was co-expressed with the indicated A3A-EGFP constructs in 293T cells, affinity purified, and analyzed by immunoblotting to detect co-purifying A3 proteins. A3A enzymes from primate species that encode *A3B* are shown in blue and A3A enzymes from primate species that lack *A3B* are shown in red.

(**B**) Schematic diagram depicting the *APOBEC3* loci from the indicated primate species representative of Hominoidea (human, genome assembly GRCh38.p13), Cercopithecinae (rhesus macaque, genome assembly Mmul_10), African *Colobinae* (Angola colobus, genome assembly Cang.pa_1.0; Ugandan red colobus, genome assembly ASM277652v2), and Asian Colobinae (golden snub-nosed monkey, genome assembly Rrox_v1). The African *Colobinae* monkey genomes available via ENSEMBL [92] indicate that the *A3B* gene has been lost, likely after the *Colobinae* subfamily split into the African and Asian tribes (10 to 14 million years ago) because its closest relatives, the Asian *Colobinae* (e.g., Golden snub-nosed monkey), still encode *A3B*. *APOBEC3* domains are colored according to Z domain classification (gold = Z1, blue = Z2, red = Z3).

(**C**) Chronogram of primates depicting the loss of *A3B* following the split of the *Colobinae* subfamily into the African and Asian tribes (indicated by delta symbol). Primate groups shown in blue are composed of species that encode *A3B* and primates groups shown in red lack *A3B*. Primate phylogeny and divergence intervals were obtained from timetree.org [85]. Mya = million years ago.

1 **Figure 8—figure supplement 4**

A

```
                                   203                L1                                   L3
Human (NM_004900.5)          MDPDTFTFNF NNDPLVLRRR QTYLCYEVER LDNGTWVLMD QHMGFLCNEA KNLLCGFYGR HAELRFLDLV
Gorilla (XM_031005506.1)     ...E...... .........H .......... .......... R......... .......... ..........
Bonobo (NM_001301766.1)      .......... .........H .......... .......... .......... .......... ..........
Chimpanzee (NM_001289639.1)  .......... .........H .......... .......... .......... .......... ..........
Orangutan (XM_009234378.2)   .......... ....F....H .........H .......K.. ..R.S.H.Q. R.P.Y.LD.. .......G.L
Gibbon (NM_001345972.1)      .......... ......FG.H .......... .......K.. ..R...H.Q. .DP.Y..D.. .T..C..G.I
Green monkey (NM_001345954.1) .......I.. ...LS..G.. .......... .......P.. ..W....Q. R.P.H.VDSC .V..C..SQ.
Sooty mangabey (XM_012032024.1) .......... ...LS..G.H .........H R......P.. ..W....Q. R.P.H.VDSC ....C...Q.
Drill (XM_011978846.1)       .......... ...LS..G.H .......... .......P.. ..W....Q. R.P.H.VDSC ....C...Q.
Baboon (NM_001346001.1)      .......S.. ...LS.RG.H .........H .......P.. ..W....Q. ....HDD..C ....C..GW.
Rhesus mac. (NM_001246230.2) .......S.. ...VS.RG.H .......... .......P.. ..W....Q. ..V.R.D..C ....C..GW.
Cynomolgus mac. (XM_005567246.2) .......S.. ...LS..G.H .......... .......P.. ..W....Q. ..V.R.D..C ....C..GW.
Pig-tailed mac. (KX583652.1) .......S.L ...LS.RG.H .......... .......P.. ..W....Q. ..VPR.D..C .V..C..GK.
Snub-nosed monkey (XM_030915811.1) ....M..... ...LSI.G.H .........H .......P.. ..W....Q. ..P.H.VD.C ....C...Q.
Ancestral                    .......... .......G.H .......... .......P.. ..R....Q. ..P.Y..D.. ....C..G..

                                   273                                                     L7
Human (NM_004900.5)          PSLQLDPAQI YRVTWFISWS PCFSWGCAGE VRAFLQENTH VRLRIFAARI YDYDPLYKEA LQMLRDAGAQ
Gorilla (XM_031005506.1)     .......... .......... .........Q .......... ...H...... .......... ..........
Bonobo (NM_001301766.1)      .......... .......... .........Q .......... .......... .......... ..........
Chimpanzee (NM_001289639.1)  .......... .......... .........Q .......... .......... .......... ..........
Orangutan (XM_009234378.2)   .YW....... .......... ........RQ .......... .......... .......... ..........
Gibbon (NM_001345972.1)      .YW...L..T .......... ........EQ .......... M......... .......... .....G....
Green monkey (XM_001345954.1) S.W......T .......... ....G...EQ .......... ...C...... .N.....Q.. .R.......
Sooty mangabey (XM_012032024.1) S.W......T .......... ........EQ .H........ .......... .......Q.. .RT.Q.....
Drill (XM_011978846.1)       S.W......T .......... ........EQ .H........ .......... .......Q.. .RT.......
Baboon (NM_001346001.1)      ..W......M .......... ....G...EQ .......... .......... .......Q.. .R.......
Rhesus mac. (NM_001246230.2) ..W......T .......... ........EQ .......... ...H...... .......Q.. .RT.......
Cynomolgus mac. (XM_005567246.2) ..W......T .......... ..V.....EQ .......... ...H...... .......Q.. .RT.......
Pig-tailed mac. (KX583652.1) ..W......T .......... ........EQ ........K. .......... .......Q.. .RT......E
Snub-nosed monkey (XM_030915811.1) ..W......M H......... ........EQ .......... .......... .......Q.. .RT.......
Ancestral                    ..W......T .......... ........EQ .......... .......... .......... ..........

                                   343
Human (NM_004900.5)          VSIMTYDEFE YCWDTFVYRQ GCPFQPWDGL EEHSQALSGR LRAILQNQGN
Gorilla (XM_031005506.1)     .........K H.......C. .......... .......... .Q........
Bonobo (NM_001301766.1)      .......... .......... .......... .......... ..........
Chimpanzee (NM_001289639.1)  .......... .......... .......... .......... ..........
Orangutan (XM_009234378.2)   .......... ...N...DH. .......... .........K .Q...L....
Gibbon (NM_001345972.1)      .......... H......DH. .R........ .......... .Q...L....
Green monkey (NM_001345954.1) ......E... .......D.. .......... D......... ..........
Sooty mangabey (XM_012032024.1) ......E... .......D.. .......... D......... .Q...-...
Drill (XM_011978846.1)       .F....K... .......D.. .R........ D......... .Q........
Baboon (NM_001346001.1)      ......E... .......D.. .......... D......... ..........
Rhesus mac. (NM_001246230.2) .......... .......D.. .R........ D......... ..........
Cynomolgus mac. (XM_005567246.2) .......... .......D.. .R........ D......... ..........
Pig-tailed mac. (KX583652.1) ......E... .......D.. .R........ D......... ..........
Snub-nosed monkey (XM_030915811.1) .F....E... .......D.. .......... D......... ..........
Ancestral                    .......... .......D.. .......... .......... ..........
```

B

Site	p-value	Most common codon substitution	Site	p-value	Most common codon substitution
216	0.099	CCT>CTT,CTT>GTT	272	0.0792	GTT>ATT,GTT>CTT
219	0.0693	CGT>CTT,CTT>CGT,CTT>TTT	273	0.0891	CCT>TCT
220	0.0891	CGA>GGA	296	0.0099	TCC>AGC
233	0.0693	CTG>CGG	304	0.099	CGT>CAT
240	0.0594	CCG>CCA,CTG>AAG,CTG>CCG	305	0.0198	GCG>GCA,GCG>TCG
249	0.0099	TGC>CAC	334	0.099	CAA>CGA
255	0.0594	CCT>CTT,CTT>GTT	337	0.099	CGG>CAG
257	0.0891	CAT>CGT,TAT>CAT,TAT>TGT	364	0.0693	TGT>CGT
259	0.0297	GTT>GAT,TTT>GTT,TTT>TTG	388	0.0891	CAG>CTG
271	0.0099	CAG>TGG,CTG>CAG,CTG>TTG,TGG>AAG			

2

1 **Figure 8—figure supplement 4**

2 (**A**) Amino acid sequence alignment of the C-terminal catalytic domain of the indicated primate

3 A3B proteins with sites under diversifying positive selection identified using MEME [46] with p

4 < 0.1 highlighted in red. L1, L3, and L7 regions are colored with blue boxes.

5 (**B**) Summary of MEME analysis results showing the p-values in dN/dS ratio test and the most

6 common codon substitutions at each codon site.

7