7848–7865 *Nucleic Acids Research, 2016, Vol. 44, No. 16* *Published online 20 July 2016*
doi: 10.1093/nar/gkw653

APOBEC3 proteins can copackage and comutate HIV-1 genomes

Belete A. Desimmie[1,†], Ryan C. Burdick[1,†], Taisuke Izumi[1,†], Hibiki Doi[1], Wei Shao[2], W. Gregory Alvord[3], Kei Sato[4,5], Yoshio Koyanagi[4], Sara Jones[6], Eleanor Wilson[2], Shawn Hill[2], Frank Maldarelli[2], Wei-Shau Hu[7] and Vinay K. Pathak[1,*]

[1]Viral Mutation Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA, [2]Clinical Retrovirology Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA, [3]Statistical Consulting, Data Management Services, Inc., Frederick, MD 21702, USA, [4]Institute of Virus Research, Kyoto University, Kyoto, 606-8057, Japan, [5]CREST, Japan Science and Technology Agency, Saitama, 332-0012, Japan, [6]Leidos Biomedical Research, Inc., Bethesda, MD 20892, USA and [7]Viral Recombination Section, HIV Dynamics and Replication Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

Received June 17, 2016; Revised July 8, 2016; Accepted July 11, 2016

ABSTRACT

Although APOBEC3 cytidine deaminases A3G, A3F, A3D and A3H are packaged into virions and inhibit viral replication by inducing G-to-A hypermutation, it is not known whether they are copackaged and whether they can act additively or synergistically to inhibit HIV-1 replication. Here, we showed that APOBEC3 proteins can be copackaged by visualization of fluorescently-tagged APOBEC3 proteins using single-virion fluorescence microscopy. We further determined that viruses produced in the presence of A3G + A3F and A3G + A3H, exhibited extensive comutation of viral cDNA, as determined by the frequency of G-to-A mutations in the proviral genomes in the contexts of A3G (GG-to-AG) and A3D, A3F or A3H (GA-to-AA) edited sites. The copackaging of A3G + A3F and A3G + A3H resulted in an additive increase and a modest synergistic increase (1.8-fold) in the frequency of GA-to-AA mutations, respectively. We also identified distinct editing site trinucleotide sequence contexts for each APOBEC3 protein and used them to show that hypermutation of proviral DNAs from seven patients was induced by A3G, A3F (or A3H), A3D and A3G + A3F (or A3H). These results indicate that APOBEC3 proteins can be copackaged and can comutate the same genomes, and can cooperate to inhibit HIV replication.

INTRODUCTION

During the last decade, numerous host restriction factors have been identified that inhibit the replication of HIV-1 and other viruses to varying degrees (1–4). Among the restriction factors reported thus far, human apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3 (APOBEC3) cytidine deaminases are among the most potent and well-characterized HIV restriction factors. The APOBEC3 superfamily consists of seven members (A3A, A3B, A3C, A3D, A3F, A3G and A3H); A3D, A3F, A3G and certain haplotypes of A3H (II, V and VII) have been shown to inhibit HIV replication (5–9). APOBEC3 proteins have specificity for single-stranded DNA and deaminate cytidines in the viral minus-strand DNA, which results in extensive G-to-A hypermutation of the viral genome during reverse transcription. In addition to the cytidine deaminase-dependent inhibition of viral replication, the APOBEC3 proteins have been shown to inhibit viral replication by inhibiting viral DNA synthesis and integration of the viral DNA into the host genome (for a recent review see Ref. (3)). The restriction activity of APOBEC3 proteins requires their incorporation into virions (7,10,11). However, lentiviruses such as HIV-1 and HIV-2 express the accessory protein viral infectivity factor (Vif), which can bind to some of the APOBEC3 proteins (A3C, A3D, A3F, A3G and A3H) and mediate their polyubiquitination and proteasomal degra-

*To whom correspondence should be addressed. Tel: +1 11 301 846 1710; Fax: +1 11 301 846 6013; Email: pathakv@mail.nih.gov
†These authors contributed equally to the paper as first authors.
Present address: Eleanor Wilson, Institute of Human Virology at the University of Maryland School of Medicine, Baltimore, MD 21201, USA.
Disclaimer: The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

dation (5,7,9,11–16). When Vif is absent or defective, the APOBEC3 proteins can be packaged into the assembling nascent virions and exert extensive cytidine deamination in the minus-strand DNA of the viral genome, most often resulting in lethal hypermutation of the viral DNA.

APOBEC3 genes have been shown to be induced by interferon (IFN) in macrophages, dendritic cells, resting $CD4^+$ T cells but not in activated $CD4^+$ T cells (17–21). A3D, A3F, A3G and A3H (haplotypes II, V and VII) have each been shown to individually inhibit HIV-1 replication, to our knowledge, there are no studies that have directly investigated the potential for different APOBEC3 proteins to copackage and to comutate the same viral genomes. A3G prefers 5′-GG editing sites and the other APOBEC3 proteins prefer 5′-GA editing sites (3); therefore, a high frequency of mutations in both GG and GA contexts in the same genome can be employed to identify copackaging of functional APOBEC3 proteins. However, a previous study analyzed nearly 100 full-length HIV genome sequences classified as 'hypermutated' viral genomes for co-existence of signature A3G- and A3F-induced G-to-A mutations by analyzing the GG and GA dinucleotide motifs of the edited sites and concluded that they rarely comutate the same genome (22). As comutation was rarely observed, it was concluded that A3G and A3F (or other A3F-like proteins) are not copackaged into the same virion. Alternatively, if they are copackaged, their copackaging does not result in comutation because only one of the APOBEC3 proteins hypermutates the viral genome irrespective of the presence of the other APOBEC3 protein. It was also suggested that A3G and A3F share a similar virion-incorporation mechanism and compete for packaging; however, most studies of virion incorporation have focused on A3G packaging and very few studies have examined virion incorporation of A3F or the other APOBEC3 proteins (23–27).

The underlying mechanisms by which APOBEC3 proteins are packaged into HIV-1 nascent virions are not fully understood and different mechanisms have been proposed. Previously, we and others have investigated the mechanism by which A3G is packaged into virions and have concluded that interactions of A3G with viral or non-viral RNAs are essential for virion incorporation (25,28–32). Other studies have proposed that APOBEC3 proteins interact with HIV-1 nucleocapsid domain of Gag in an RNA-dependent manner and that this interaction is essential for virion incorporation (30,33,34). It was also proposed that interactions with 7SL RNA and/or other small RNAs are critical for virion incorporation (25,32,35), while other studies found that interactions with 7SL RNA are dispensable for virion incorporation (27–28,36,37). Furthermore, Apolonia *et al.* demonstrated that the promiscuity of A3G and A3F to non-specifically bind to a variety of RNAs ensures their effective encapsidation into HIV-1 particles (36). In addition to their high affinity to RNAs (36,38,39), APOBEC3 proteins are also known to associate with cellular ribonucleoprotein complexes in the cytosol (37,40,41), suggesting that some unidentified mechanisms may govern the specificity of their virion incorporation.

APOBEC3 proteins are expressed in human $CD4^+$ T cells (17–19,42,43) and have been shown to hypermutate HIV-1 DNA in infected cells (42,44–50); however, it is not known whether different APOBEC3 proteins can copackage into the same virions. Although not required for their antiviral action, A3G and A3F were shown to co-localize in mRNA processing (P) bodies and hetero-oligomerize through an RNA-dependent interaction (9,37,51), suggesting that they may bind the same RNA and incorporate into the same virions. Furthermore, their distinct editing site dinucleotide preferences suggest that they may not compete for targets in the genome and could induce comutation of the same genomes. In addition, *in vitro* biochemical studies have demonstrated that APOBEC3 proteins exhibit different mechanisms of movement on substrate nucleic acid template, which perhaps contribute to their editing site selection and specificity; such studies were done for A3G, A3F as well as A3H haplotype II and haplotype V (52,53).

Here, we determined the potential of APOBEC3 proteins to copackage and their ability to comutate the same viral genomes. We used single-virion analysis to unambiguously show that A3G can efficiently copackage into the same HIV-1 particles with A3F, A3D or A3H haplotype II (hereafter referred to as A3H). We analyzed the trinucleotide sequence contexts in which A3G, A3F, A3D and A3H induced hypermutation in cell culture-based experiments as well as in proviral DNA isolated from HIV-1 infected patients. Our results suggest that A3G can be copackaged with the other APOBEC3 proteins (A3F, A3D or A3H) into the same virions and can comutate the same viral genome in single-cycle infection assays as well as during natural HIV-1 infection *in vivo*.

MATERIALS AND METHODS

Cell culture, transfections and production of VSV-G-pseudotyped HIV-1 containing APOBEC3 proteins

293T and CEM-SS cell lines were obtained from the American Type Culture Collection and maintained in Dulbecco's modified Eagle's medium and RPMI 1640 medium (Corning Cellgro), respectively. Both media were supplemented to contain 10% fetal calf serum (Hyclone), 100 IU/ml penicillin and 100 μg/ml streptomycin (GIBCO).

All viruses were prepared by using a previously described HIV-1-based vector pHDV-EGFP (54), which expresses Gag-Pol and eGFP from the HIV-1 LTR promoter. phCMV-G plasmid, which expresses vesicular stomatitis virus glycoprotein (VSV-G), was used to pseudotype the viral particles (55). Venus yellow fluorescent protein (YFP)-tagged APOBEC3G (A3G-YFP) and/or mCherry-tagged A3F (A3F-mCH), mCherry-tagged A3D (A3D-mCH) and mCherry-tagged Hap II A3H expression plasmids were co-transfected during virus production to label the particles as described previously (37). Briefly, 4×10^6 cells 293T cells were seeded and co-transfected the next day with pHDV-EGFP (10 μg), phCMV-G (0.5 μg) and pA3G-YFP (5 μg), pA3F-mCH (5 μg), pA3D-mCH (5 μg), pA3H-mCH (5 μg) or pA3G-YFP (5 μg) and pA3F-mCH (5 μg), pA3D-mCH (5 μg) or pA3H-mCH (5 μg) using polyethylenimine as previously described (56). For visualization of virions, non-infectious Gag-cerulean fluorescent protein (Gag-CeFP)-labeled viral particles were produced using the same cocktails of plasmids described above except pHDV-EGFP was replaced with a mixture of pGag-CeFP-BglSL (5 μg)

and pGag-BglSL (5 μg) (57,58). The culture supernatant containing virus was harvested at 24-h (for infection and proviral DNA hypermutation analysis) or 18-h (for single virion analysis) post-transfection, filtered through 0.45-μm filters and concentrated by ultracentrifugation through a 20% sucrose cushion at 25 000 rpm for 1.5 h at 4°C as described previously (37,58). The virus pellets were resuspended in culture medium and kept at −80°C until use.

APOBEC3 protein titration and viral infectivity assays

Human embryonic kidney 293T cells were seeded at 8 × 10^5 cells per well in 6-well plates. The transfection procedure was the same as above and the plasmid mixes contain 1 μg of the HIV-1 vector pHDV-EGFP and 0.2 μg of VSV-G plasmid phCMV-G. To this mix we added A3G-YFP, A3F-mCH, A3D-mCH and/or A3H-mCH. All APOBEC3 proteins were titrated serially by 2-fold (from 15.6 to 500 ng). For some combination experiments, we kept A3G-YFP at 62.5 ng and added different amounts of three point titration of the other APOBEC3 proteins. In all cases, to maintain equal amounts of DNA, pcDNA3.1noMCS was used when needed. After 48 h, the virus-containing supernatant was clarified through 0.45-μm filters and virus preparations were kept at −80°C until use.

To determine the infectivity of viruses produced in the presence of the different APOBEC3 proteins, TZM-bl cells were seeded in 96-well plates (4 × 10^3 cells per well) a day prior to infection. Next day, we infected TZM-bl cells in triplicates with viruses normalized for p24 (XpressBio) and 48-h later luciferase activity was determined using britelite™ plus (PerkinElmer) following the instructions of the manufacturer and measured using a LUMIstar Galaxy luminometer or 1450 MicroBeta JET (PerkinElmer).

Western blot analysis

Virus samples and cell lysates were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and immunoblot analysis. Virus samples input were normalized for p24 by ELISA. While A3G-YFP was detected using anti-GFP (Sigma), A3F-mCH, A3D-mCH and A3H-mCH were detected using anti-mCherry antibodies (Abcam). The precursor p55 Gag and p24 CA were detected using a mouse anti-HIV-1 p24 CA monoclonal antibody (kindly provided by Michael H. Malim, AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH). Tubulin, a loading control for cell lysates, was detected with mouse anti-tubulin antibody (Sigma). The primary antibodies were probed using an IRDye 800CW- or IRDye 680-labeled secondary antibodies (Li-COR). The signal intensities of the protein bands were calculated using the Odyssey system (Li-COR).

Confocal microscopy

Human embryonic kidney 293T cells (20 000 cells per well) were seeded onto ibiTreat 8-well μ-slides (Ibidi) and co-transfected with A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH expression plasmids using the Lipofectamine™2000 transfection reagent (Invitrogen). The cells were fixed 18-h post-transfection with 4.0% paraformaldehyde and cellular DNA was counterstained with DAPI (Sigma). The transfected cells were visualized using a LSM780 laser-scanning confocal microscope (Zeiss).

Single-virion analysis using fluorescence microscopy

Single-virion analysis was performed as described previously (37,57,58) and used to directly visualize virion incorporation of A3G-YFP, A3F-mCH, A3D-mCH or A3H-mCH (37,58). Concentrated virus supernatant (0.4 μl) was briefly mixed with 200 μl phosphate buffered saline and then centrifuged onto ibiTreat μ-slides (Ibidi) at 1200 × *g* for 1 h. Images of virus-like particles (VLPs) were acquired using epifluorescence microscopy (Nikon). CeFP was used to identify VLPs containing Gag, and YFP and mCH were used to detect virion incorporated A3G-YFP and A3F-mCH, A3D-mCH or A3H-mCH proteins, respectively (37,57,58). The number, positions and integrated intensities of the spots were quantified for each channel using Localize software as previously described (37,57–59). The positions of the spots were also used to determine colocalization; spots were considered colocalized if the centers of the spots were within three pixels of each other. Merged and pseudocolored images were generated using NIS Elements (Nikon). A3G, A3F, A3D and A3H virion incorporation efficiency was calculated by determining the percentage of $CeFP^+$ particles that contained the YFP and/or mCH signals, respectively, by using an in-house MATLAB program (MathWorks).

Proviral DNA sequencing and hypermutation analysis

CEM-SS cells (3 × 10^5 cells/well in 24-well plates) were infected with 30 ng of p24 CA equivalent HDV-EGFP virus preparations, as determined by ELISA. Cells were collected at 24-h post-infection and genomic DNA was extracted using QIAamp Blood DNA Mini Kit following the manufacturer's instruction (QIAGEN). Approximately 2-kb of the *pol* (NL4-3 2258–4343 bp) region encompassing the RT reading frame was amplified using a forward primer HIV-07 ([NL4-3 #2258-2280]: 5′-GATC ACTCTTTGGCAGCGACCCC-3′) and a reverse primer HIV-11b [NL4-3 #4343-4317]: 5′-GGCTACTATTTC TT TTGCTACTACAGG-3′) (see the schematic in Figure 3A). The polymerase chain reaction (PCR) primers were designed to anneal in regions with the fewest possible putative A3F and A3G deamination sites to avoid potential primer mismatch due to A3-induced mutagenesis. The PCR products were purified using QIAquick PCR Purification Kit (QIAGEN) and cloned into pCR2.1 vector with TA Cloning Kits (Life technologies). The insert was then sequenced using the following overlapping primers: HIV-08N-for (NL4-3 #2553-2590): 5′-ATTAGTCCTATT GAGACTGTACCAGTAAAATTAAAGCC-3′; HIV-08-rev (#3043-3018): 5′-GTCATGCTACACTGGAATATTG CTGG-3′; HIV-10-for (#3296-3321): 5′-GGACAGCTGG ACTGTCAATGACATAC-3′; and HIV-11-rev (#4188-4162): 5′-GTTCATTTCCTCCAATTCCTTTGTGTG-3′.

Study participants

HIV infected Individuals ($N = 7$) were enrolled in studies at the NIAID/CCMD clinic at the NIH Clinical Center in Bethesda MD and were sampled in the period 1994–2004. All participants were male with median age 40.8 years undergoing antiretroviral therapy with persistent viremia (see Supplementary Table S1 for demographic information); 4/7 had diagnosis of AIDS. Plasma and peripheral blood mononuclear cells (PBMCs) were obtained during protocol visits by phlebotomy.

Ethics statement

All participants in this study were enrolled in clinical protocols approved by the NIAID Institutional Review Board (FWA00005897) administered at the NIH Clinical Center in Bethesda, MD, USA. Individuals underwent an informed consent process and provided written consent for participation in the research described here.

APOBEC3H genotyping and haplotype determination

Genomic DNA was extracted from PBMCs and A3H exons were amplified with PCR Supermix High Fidelity (Invitrogen) using primers and PCR genotyping conditions for previously identified A3H amino acid polymorphisms (ΔN15, rs139292; R18L, rs139293; G105R, rs139297; D121K, rs139299 and rs139298; E178D, rs139302) described by OhAinle *et al.* (60). PCR products were gel eluted and directly sequenced as well as cloned into TA cloning vector (Invitrogen), in which case multiple clones (4–8 clones) were sequenced.

Viral RNA and proviral DNA sequence analysis

Plasma viral RNA and PBMC proviral DNA were analyzed using single-genome sequencing (SGS) covering ∼1100-bp *pro-pol* region as previously described (61,62). For each patient, alignments were constructed for both the plasma derived and plasma derived and PBMC DNA derived samples (MEGA 5.0). A patient-specific consensus was obtained from the plasma alignment, and used as a standard sequence to detect hypermutation in HIV from PBMC derived DNA. Each PBMC derived HIV sequence was analyzed for the presence of GG and GA motifs present in the plasma consensus. Patient cellular DNA sequences were considered hypermutated if the number of mutations was significantly above genetic variation in the plasma viral RNA population for each patient. A sequence was considered to be hypermutated in the GG context if it contained at least 5–7 mutations compared to the consensus plasma RNA sequence (average 0–1 mutations/clone). A sequence was considered hypermutated in the GA context if it contained at least 7–10 mutations compared to the consensus plasma RNA sequence (average 1–3 mutations/clone).

Statistical analysis

Data were analyzed with paired and Welch's unpaired *t*-tests, generalized linear models and loglinear modeling techniques. The numbers of G-to-A mutations were tested for homogeneity (independence), and were further partitioned (as necessary) into orthogonal components such that independent inferences could be drawn for each component involved in the partitioning (63). For some analyses, *post hoc* Fisher's Exact tests were performed on 2×2 subtables that were extracted from 4×2 tables and the *P*-values for these tests were extracted. The *P*-values for these *post hoc* tests were adjusted using the Bonferroni correction for multiple comparisons. Calculations were performed with the R Statistical Language and Environment (R Core Team) (https://www.r-project.org) (64). Probability values <0.05 ($P < 0.05$) were considered significant.

RESULTS

Virion incorporation of APOBEC3 proteins into Vif-deficient HIV-1 particles

To determine whether APOBEC3 proteins can be copackaged into the same virus particles, we constructed expression vectors in which A3G was tagged at its C-terminus with venus YFP (A3G-YFP); we also constructed vectors in which A3F, A3D and A3H were tagged at their C-termini with monomeric cherry fluorescent protein (A3F-mCH, A3D-mCH and A3H-mCH, respectively) (Figure 1A). A3G-YFP, A3F-mCH, A3D-mCH or A3H-mCH were co-expressed alone or in combination in 293T cells with HDV-EGFP, an HIV-1 vector that expresses Gag-Pol and is defective in expression of Vif and Env (Δ*vif* Δ*env*). Virus particles were harvested 24-h later, and the virus producer cell lysates and viral lysates were analyzed by western blotting analysis (Figure 1B and C). All APOBEC3 proteins were readily detected in cells transfected with A3G-YFP only, A3F-mCH only, A3D-mCH only, A3H-mCH only, A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH or A3G-YFP + A3H-mCH (Figure 1B). Quantitation of the western blots indicated that similar levels of A3G were expressed in cells transfected with A3G-YFP alone and in combination with the other APOBEC3 proteins, and similar levels of A3F, A3D and A3H were expressed in cells transfected only with A3F-mCH, A3D-mCH or A3H-mCH, or in combination with A3G-YFP. Moreover, all APOBEC3 proteins exhibited mostly diffuse distribution in the cytosol of the producer cells when expressed alone or in combination (Supplementary Figure S1). As previously observed by us and others, we also observed accumulation of all APOBEC3 proteins in cytosolic puncta that were previously shown to be RNA processing bodies (P bodies (37,51,65).

We next determined the relative amounts of A3G, A3F, A3D and A3H packaged proteins into purified HIV-1 particles based on the respective APOBEC3 proteins incorporated into virions when expressed alone (Figure 1C); the amounts of viruses used for western blots were normalized by p24 CA amounts using ELISA. Virion incorporation of A3G-YFP was similar in cells transfected with A3G-YFP only or in combination with A3F-mCH, A3D-mCH and A3H-mCH, respectively. The relative virion incorporation of A3F-mCH, A3D-mCH and A3H-mCH was also similar in viruses produced in cells transfected with or without A3G-YFP. Of note, the packaging efficiencies of A3D-mCH only and A3H-mCH only were 5- and 10-fold less, respectively, compared to A3F-mCH only; pack-

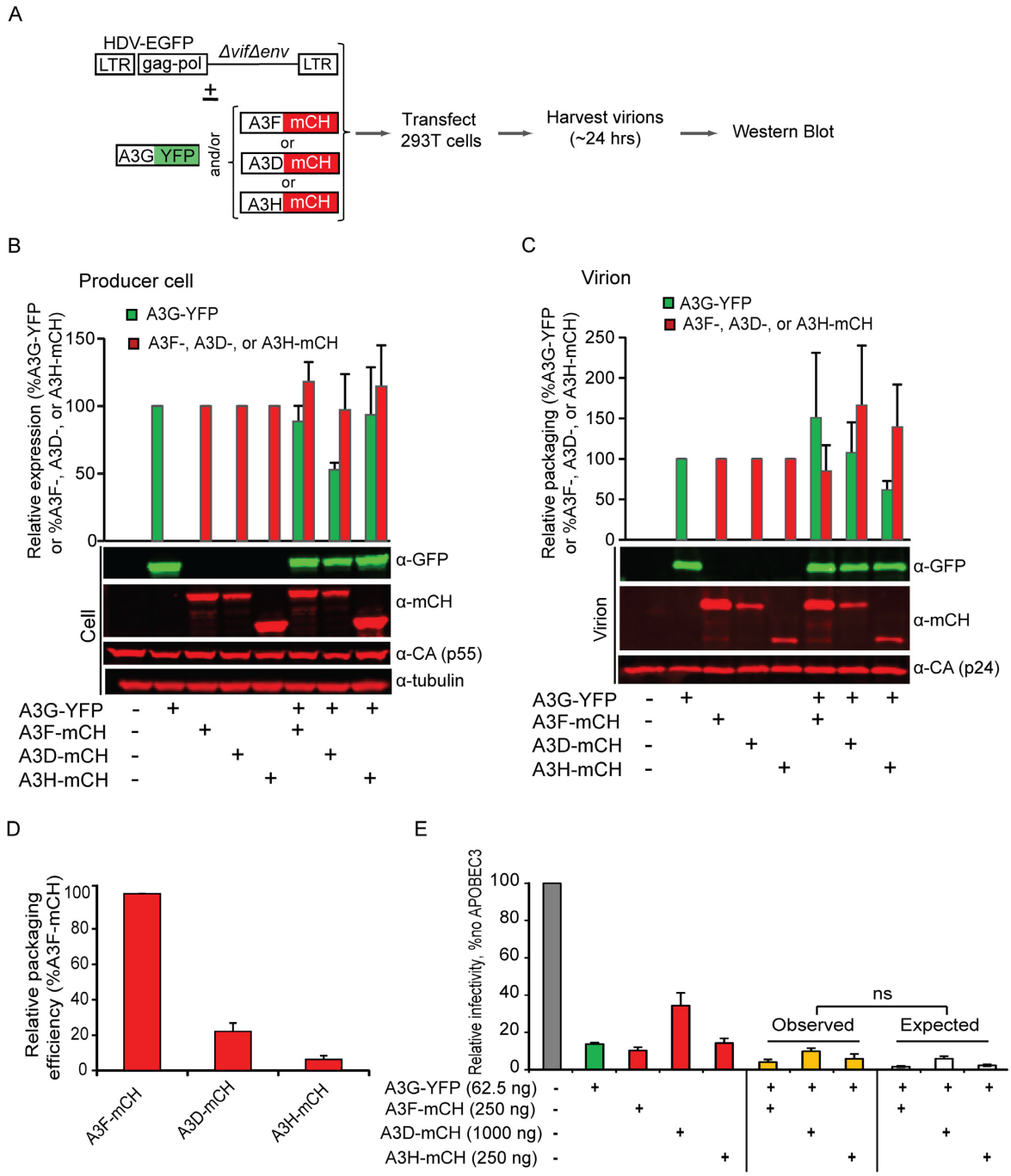

Figure 1. Analysis of copackaging of multiple APOBEC3 proteins into *Δvif* HIV-1 particles. (**A**) Schematics of plasmids used to produce viral particles and western blot analysis of purified viral particles and producer cells lysates. (**B**) Representative immunoblots and quantitation of the expression levels of A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH proteins in producer cell lysates were analyzed by western blotting; p55 Gag and tubulin bands were shown as control. The bar graphs indicate the average expression of A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH in the combination experiments relative to A3G-YFP only, A3F-mCH only, A3D-mCH only or A3H-mCH only, respectively, which were set to 100%. Quantitation was done after normalization for tubulin measured using the Odyssey system. (**C**) Representative immunoblots and quantitation of the A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH proteins in virus lysates. The level of A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH proteins after normalization for p24 CA was determined using the Odyssey system relative to A3G-YFP only, A3F-mCH only, A3D-mCH only or A3H-mCH only, respectively. (**D**) Analysis of the relative packaging efficiency of A3D-mCH and A3H-mCH compared to A3F-mCH (set to 100%). (**E**) Relative infectivity of viruses produced in the presence of the four human APOBEC3 proteins either individually or in combination with A3G compared to the no APOBEC3 condition (set to 100%). Single-cycle HDV-EGFP virus was produced in the presence or absence of human APOBEC3. The A3G concentration was determined after titration and 62.5 ng of plasmid, which led to ∼80% drop in infectivity, was used in these experiments. The expected inhibition of infectivity (white bars) was calculated by multiplying the inhibition by the individual protein and then divided by 100. All quantitative data are from three independent experiments; n.s., no significant differences.

aging efficiency was calculated as normalized band intensity of virion-associated protein relative to normalized cellular expression of the respective protein (Figure 1D). Taken together, these results suggest that co-expressed A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH and A3G-YFP + A3H-mCH can be packaged into HIV-1 particles independent of each other.

We next assessed the infectivity of viruses produced in the presence of the four APOBEC3 proteins expressed alone or A3G-YFP co-expressed with one of the other APOBEC3 proteins. First, we carefully evaluated the antiviral activity of each APOBEC3 protein by titration of the plasmids transfected in virus producer cells. VSV-G pseudotyped *vif*-deficient HDV-EGFP single-round virus stocks were produced in the presence of escalating doses of each of the four human APOBEC3 proteins with the same fluorescent protein tags as described above and the infectivity of the viruses was determined in TZM-bl cells using p24 CA-normalized virus input. Inhibition of infectivity by 60–80% compared to the no APOBEC3 control was used to determine the amount of plasmids used in the APOBEC3 proteins combination experiments. Figure 1E shows the relative infectivity of viruses produced in the presence of each APOBEC3 protein alone or combination of A3G-YFP and one of the other three APOBEC3 proteins. We observed a significant reduction in the infectivity of virus produced in the presence of A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH or A3G-YFP + A3H-mCH compared to the respective individual proteins. Co-expression of A3G with one of the other APOBEC3 proteins in the virus producer cells resulted in additive antiviral activities, since the observed antiviral activities were not significantly different from the expected combined antiviral activities of the individual APOBEC3 proteins (the product of their individual antiviral activities; $P > 0.05$, *t*-test for comparison of observed versus expected antiviral activities).

Single-virion analysis of the copackaging efficiency of APOBEC3 proteins

The western blotting results shown in Figure 1, or other biochemical analyses, cannot determine whether A3G-YFP and A3F-mCH, A3D-mCH or A3H-mCH are copackaged into the same virions or into different populations of virus particles. To determine whether A3G is copackaged into the same virions with the other APOBEC3 proteins, we exploited the recently developed single-virion fluorescence microscopy imaging technique (57,58). We produced virus by co-transfecting plasmids expressing wild-type HIV-Gag, HIV-Gag-CeFP and A3G-YFP and/or A3F-mCH, A3D-mCH or A3H-mCH expression plasmids into 293T cells (Figure 2A). Transfection of a mixture of HIV-Gag-CeFP and unlabeled HIV-Gag expression plasmids (1:1) produces non-infectious VLPs. Previous studies have shown that >90% of the CeFP$^+$ signals colocalized with virion RNA signals and are thus bona fide viral particles (57). Although some A3G-YFP and A3F-mCH proteins are released from transfected cells in vesicles such as exosomes, we only analyzed CeFP$^+$ particles that contained YFP and/or mCH signals, and the APOBEC3 proteins in exosomes did not have any effect on our analysis.

An example of single-virion analysis is shown in Figure 2B. The results showed that when Gag-CeFP was co-transfected with A3G-YFP alone (top panels) or A3F-mCH alone (middle panels), Gag-CeFP$^+$ particles that were YFP$^+$ or mCH$^+$, respectively, could be detected. When Gag-CeFP was co-transfected with A3G-YFP + A3F-mCH (bottom panels), Gag-CeFP particles that were both YFP$^+$ and mCH$^+$ could be detected. Co-localization of Gag-CeFP with A3G-YFP + A3F-mCH indicated that both A3G-YFP and A3F-mCH were copackaged into the same HIV-1 particles.

We then analyzed the proportion of HIV particles that incorporated detectable levels of the four human APOBEC3 proteins (A3D, A3F, A3G and A3H) by quantifying ∼9000 to ∼20 000 Gag-CeFP$^+$ particles from three independent experiments for each condition (Figure 2C). The average labeling efficiencies of Gag-CeFP$^+$ virions produced in the presence of A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH were 78.3 ± 4.5%, 81.8 ± 3.8%, 50.0 ± 6.4% and 83.0 ± 4.7%, respectively. When A3G-YFP and A3F-mCH were co-expressed, A3G-YFP was detected in 66.8 ± 5.8% and A3F-YFP was detected in 76.5 ± 5.7% of the particles; similarly, when A3G-YFP and A3D-mCH were co-expressed, A3G-YFP was detected in 73.6 ± 4.2% and A3D-mCH was detected in 42.0 ± 7.6%, and when A3G-YFP and A3H-mCH were co-expressed, A3G-YFP was detected in 72.8 ± 2.5% and A3H-mCH was detected in 79.7 ± 7.3% of virions. We found that in the combination experiments ∼58%, ∼37% and ∼64% of HIV-1 particles contained A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH and A3G-YFP + A3H-mCH, respectively. The proportions of virions containing APOBEC3 proteins when single APOBEC3 proteins were expressed were not significantly different from the proportions of particles containing detectable levels of these enzymes in the combination experiments ($P > 0.05$ for all comparisons). This observation indicated that, contrary to a previous report (66), APOBEC3 proteins did not have much (if any) influence on the packaging efficiency of the other APOBEC3 protein.

To quantify the amounts of APOBEC3 proteins that were packaged when they were expressed alone or in combination with other APOBEC3 proteins, we compared the fluorescence intensities of YFP and mCH in single and dual-labeled virions that were produced from cells expressing a single APOBEC3 protein or from cells expressing A3G-YFP and one of the other APOBEC3 proteins (Figure 2D). For this analysis, the A3G-YFP only and A3F-mCH only were set to 100% and the fluorescence intensities of the others were plotted relative to the respective fluorescent protein-tagged APOBEC3 protein. The fluorescence intensities of all APOBEC3 proteins were comparable in both single- and dual-labeled HIV-1 particles, further indicating that copackaging of A3G-YFP with the other APOBEC3 proteins did not influence the amounts of the A3G-YFP or other APOBEC3 proteins copackaged into the same particles ($P > 0.05$; t-test). Interestingly, the proportion of virions containing detectable levels A3D-mCH is ∼2-fold less compared to the other APOBEC3 proteins (Figure 2C); furthermore, the fluorescence intensity analysis suggests that the virions incorporating A3D-mCH contained about half the number of A3D-mCH molecules compared to the other

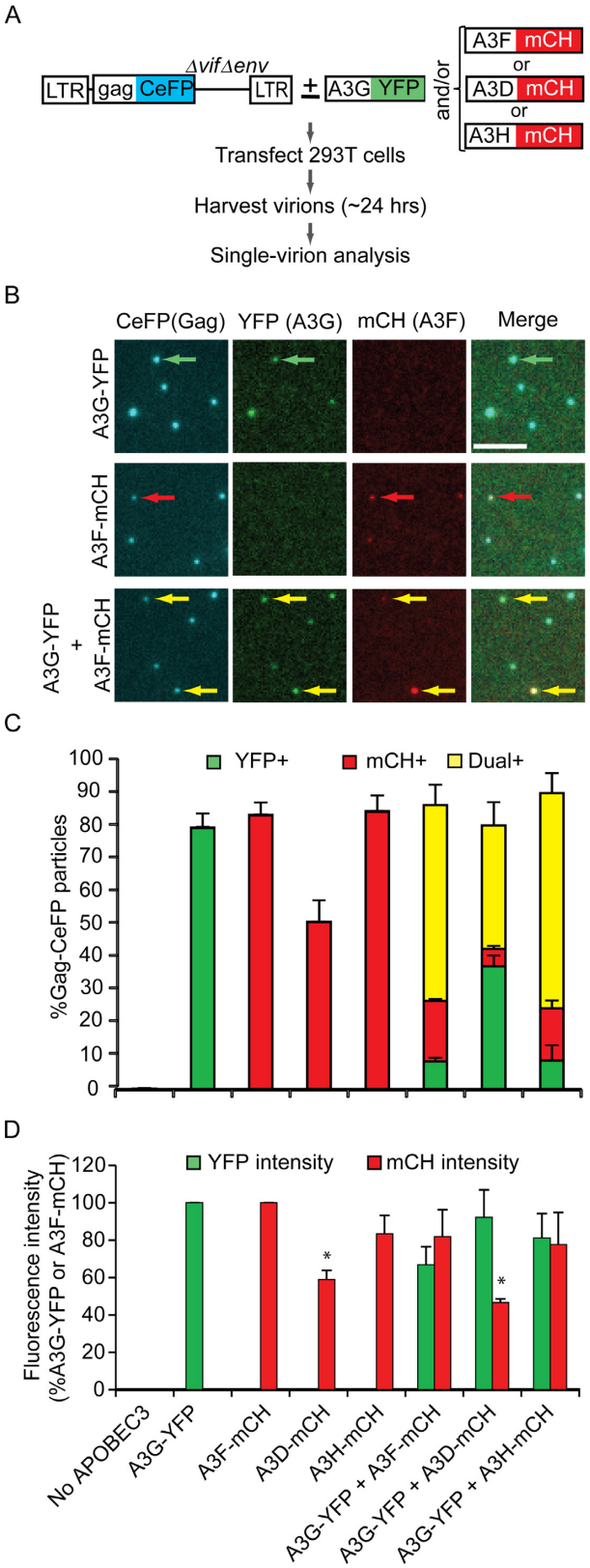

Figure 2. Determination of A3G and A3F, A3D or A3H copackaging using single-virion analysis. (**A**) Protocol for production of virus particles and single-virion analysis for A3G-YFP, and A3F-mCH, A3D-mCH or A3H-mCH copackaging. Plasmids expressing Gag and Gag-CeFP (1:1 ratio) were co-transfected with plasmids expressing A3G-YFP and/or A3F-, A3D- or A3H-mCH into 293T cells, virus particles were harvested 24 h after transfection and single-virion analysis was performed using fluorescence microscopy. (**B**) Representative images of virus particles labeled with CeFP (Gag), YFP (A3G) and/or mCH (A3F) are shown. The top row of panels shows an example of a particle labeled with Gag-CeFP that is also labeled with A3G-YFP (green arrows). The middle row of panels shows an example of a particle labeled with Gag-CeFP and A3F-mCH (red arrows). The bottom row of the panels shows an example of a particle that is labeled with A3G-YFP and A3F-mCH (yellow arrows) (scale bar, 2 μm). (**C**) Comparison of the proportion of Gag-CeFP$^+$ particles that are A3G-YFP$^+$ (green bars) and A3F-, A3D- or A3H-mCH$^+$ (red bars) in cells transfected with A3G-YFP, A3F-mCH, A3D-mCH or A3H-mCH alone or co-transfected with A3G-YFP + A3F-, A3D- or A3H-mCH. The proportion of YFP$^+$ and/or mCH$^+$ particles was determined in three independent experiments (~9000–24 000 particles per experimental group). (**D**) Relative integrated fluorescence intensity of A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH in dual-labeled particles relative to A3G-YFP only or A3F-mCH only single-labeled particles. Error bars represent SD from three independent experiments. Welch's unpaired *t*-test was performed to determine statistical significance (*$P < 0.05$).

APOBEC3 proteins (Figure 2D; *$P < 0.05$). Both observations are likely to be due to the reduced levels of A3D-mCH expression compared to the other APOBEC3 proteins (Figure 1C). In any case, the reduced virion incorporation of A3D can at least partially explain its reduced antiviral potency that we and others have observed (Figure 1E, (42,48)). Taken together, these results show that APOBEC3 proteins can be copackaged without any apparent reduction in their packaging efficiencies.

Analysis of patterns of hypermutation in the presence of multiple APOBEC3 proteins

APOBEC3 proteins are known to edit cytidines during synthesis of the minus-strand DNA by selectively targeting 5′-CC-3′ (A3G) and 5′-TC-3′ (A3D, A3F and A3H) dinucleotide sites (7,42,60,67,68). Moreover, these members of the APOBEC3 family of proteins have been demonstrated to be expressed in the natural target cells of HIV (18,43). To study the effect of copackaged APOBEC3 proteins on comutation of the same viral genome, we used viruses produced in the presence of A3G-YFP only, A3F-mCH only, A3D-mCH only, A3H-mCH only, A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH or A3G-YFP + A3H-mCH to infect CEM-SS cells and analyzed the pattern of hypermutations in HIV proviral DNA (Figure 3; Tables 1 and 2). A 1320-nt region of HIV-1 RT was PCR-amplified from proviral DNAs 24-h post-infection, cloned into a plasmid and individual PCR products were sequenced (Figure 3A). GG-to-AG mutations (predominantly induced by A3G) and GA-to-AA mutations (predominantly induced by A3F, A3D and A3H) will be hereafter referred to as GG and GA mutations, respectively. In total, we analyzed 122–250 independent clones from three to six independent experiments performed with viruses produced from cells expressing A3G-YFP only, A3F-mCH only, A3D-mCH only, A3H-mCH only, A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH and A3G-YFP + A3H-mCH (Table 1). Because of differences in their deamination activity, the frequency of clones that were hypermutated ranged from 9% for A3D to 96% for A3G (Table 1). Almost all of the sequences an-

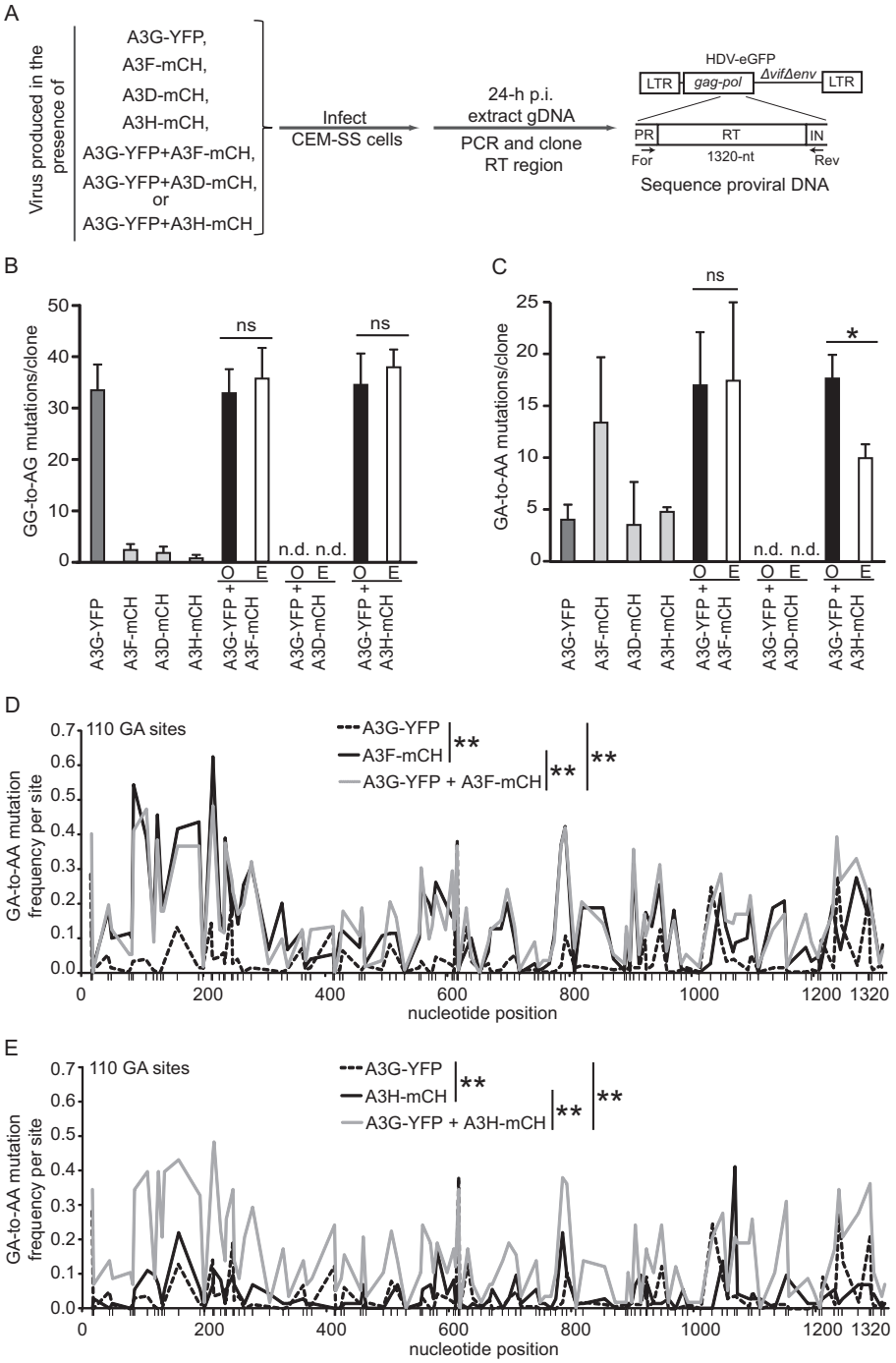

Figure 3. APOBEC3 proteins can comutate the same HIV genome. (**A**) Schematic representation of CEM-SS infection with the indicated HDV-EGFP viruses made in the presence of A3G-YFP only, A3F-mCH only, A3D-mCH only, A3H-mCH only, A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH or A3G-YFP + A3H-mCH. Genomic DNA was extracted from the infected cells 24-h post-infection (p.i.), and a 1320-bp fragment encoding reverse transcriptase (RT) was PCR-amplified using the indicated forward and reverse primers and sequenced using four overlapping primer sets. Numbers of GG-to-AG (**B**) and GA-to-AA (**C**) mutations per clone are shown for A3G (dark gray bars), A3D, A3F and A3H (light gray barsl), for the combination experiments, the calculation was done using dual-hypermutated sequences only (observed frequencies [O]; *black bars*). The A3G-YFP + A3D-mCH conditions were not determined (n.d.) because of the limited numbers of dual-hypermutated sequences. The expected frequencies (E; white bars) were calculated by adding the mutation frequencies/clone of each APOBEC3 protein from Table 1. The averages ± S.D. from three to six independent experiments are shown. Welch's unpaired *t*-test was performed to determine statistical significance (*$P < 0.005$; ns, not significant). (**D**) The distribution of GA-to-AA mutations induced by A3G-YFP only (dashed line), A3F-mCH only (black line) and A3G-YFP + A3F-mCH (light gray line), are shown relative to their locations in the 1320-nt RT region. (**E**) The distribution of GA-to-AA mutations induced by A3G-YFP only (dashed line), A3H-mCH only (black line) and A3G-YFP + A3H-mCH (light gray line). Each point represents the average frequency of mutations/site after combining the sequences from three to six independent experiments (1.0 = 100%). Black vertical tick lines on the X-axis indicate positions of 110 GA target sites (D and E). Data were analyzed with generalized linear model and loglinear modeling techniques. See text for a detailed description of the non-homogeneity statistical analysis and group comparisons (**$P < 0.0001$).

alyzed had unique G-to-A mutation patterns, confirming that they were obtained from independent proviral DNAs; the few sequences that were duplicates (<4 per group) were removed from the analysis to avoid bias resulting from amplification of some clones during PCR.

The mutations induced by A3G-YFP, A3F-mCH, A3D-mCH and A3H-mCH when expressed alone or in combination were in agreement with the expected dinucleotide sequence contexts (48,67,69) as summarized in Table 1, Figure 2B and C. A3G predominantly edited GG sites (88% of the mutations; 32.1 mutations/clone) and ∼12% of the mutations occurred at the GA sites (3.9 mutations/clone). On the other hand, A3D, A3F and A3H predominantly edited GA sites at frequencies of 3.4, 13.0 and 4.6 mutations/clone, respectively; however, they also induced some mutations at GG sites (0.6–2.2 mutations/clone) (Table 1). We then assessed the hypermutation patterns of cells infected with virions produced in the presence of A3G-YFP and one of the other APOBEC3 proteins (A3F-mCH, A3D-mCH or A3H-mCH), and as expected, the frequencies of clones that were hypermutated were >90% (Table 1).

To determine whether there were any potential interactions between copackaged APOBEC3 proteins that influence the pattern of hypermutation, we compared the observed and expected frequencies of hypermutation at GG and GA sites (Figure 3B and C). The expected frequencies of hypermutation were determined by summing the hypermutation frequencies observed when the APOBEC3 proteins were individually expressed (Figure 3B and C, white bars). Sequences that were hypermutated by A3G and another APOBEC3 protein were defined as dual-hypermutated sequences if they met the following two criteria (Table 1). First, we determined the percentage of G-to-A mutations that were at GA sites for virions produced in the presence of A3G-YFP only (12%); proviral sequences that exhibited a high percentage of GA mutations (>2 standard deviations above the mean for virions produced in the presence of A3G-YFP only) were defined as dual-hypermutated sequences. Second, the sequence context in which the GA mutations occurred (discussed in more detail below) was analyzed to verify that the GA mutations were generated by the action of A3G (GAG), A3F/A3H (GAA) or A3D (GAT). In the presence of A3G-YFP + A3F-mCH and A3G-YFP + A3H-mCH, the frequencies of dual-hypermutated proviruses were 49.6 and 48.7%, respectively (Table 1), which was in general agreement with their copackaging efficiencies of 58 and 67%, respectively (Figure 2C; yellow bars). The frequencies of dual-hypermutated clones may have been lower than the copackaging efficiencies because of deamination-independent inhibition of DNA synthesis (70–73). However, we found only 2/113 (1.8%) dual-hypermutated clones in the presence of A3G-YFP + A3D-mCH. Since both A3D and A3H can mutate GA sites with similar frequencies (Table 1), the result indicated that A3D did not significantly increase the frequency of dual-hypermutated clones because it is inefficiently incorporated into virions.

The presence of A3F or A3H in virions that also contained A3G-YFP did not alter the average mutation frequency at the GG sites (32.1, 31.6 and 33.1 mutations/clone for A3G-YFP only, A3G-YFP + A3F-mCH and A3G-YFP + A3H-mCH, respectively; Table 1 and Figure 3B; $P > 0.05$). Co-expression of A3G-YFP + A3D-mCH resulted in very few dual-hypermutated clones (2/113), which precluded any further detailed analysis of potential interactions between A3G and A3D. These data suggested that co-expression of A3F-mCH or A3H-mCH did not influence the ability of A3G-YFP to induce mutations at GG sites (Table 1, Figure 3B; $P > 0.05$).

Analysis of mutation frequencies at GA sites revealed that the presence of A3G-YFP + A3F-YFP or A3G-YFP + A3H-mCH resulted in 16.4 and 17.1 mutations per clone in the dual-hypermutated sequences, respectively, which was not significantly higher than the expected value for A3G only + A3F only samples (3.9 + 13.0 = 16.9 mutations/clone; Table 1 and Figure 3C; $P > 0.05$), but was significantly higher than the expected value for A3G only + A3H only samples (3.9 + 4.6 = 8.5 mutations/clone; Table 1 and Figure 3C; $P < 0.005$). This result suggested that the copackaging of A3G-YFP and A3H-mCH resulted in a synergistic increase in mutations at the GA sites.

Although we did not observe an increase in the frequency of mutations at GG sites when A3G-YFP was expressed alone or in combination with A3F-mCH or A3H-mCH (Table 1 and Figure 3B), potential interactions between APOBEC3 proteins could have influenced the selection of GG target sites by A3G-YFP. To address this question, we analyzed the distribution of mutations in the 1.3-kb region of *pol* which contains 78 potential GG editing sites by applying the Likelihood Ratio Chi Square statistical analysis (Supplementary Figure S2A and B). Although the edited GG sites were distributed throughout the sequence, we observed that the distribution of mutations at GG sites was not homogeneous across sites (Supplementary Figure S2A and B; $P < 0.0001$), indicating that some GG sites were preferred targets for mutation as compared to other GG sites. Moreover, when we compared the A3G-YFP only samples to A3G-YFP + A3F-mCH or A3G-YFP + A3H-mCH, we found that the distribution of GG edited sites was homogenous and overlapped in all the three groups, indicating that copackaging of A3F-mCH or A3H-mCH did not significantly influence the A3G deamination activity at the GG sites. Since A3F-mCH and A3H-mCH induced very few mutations at GG sites (2.2 and 0.6 mutations/clone, respectively, compared to 32.1 mutations/clone by A3G; Table 1), their contribution to mutations at GG sites did not have a significant influence on the overall distribution of mutations at GG sites.

We performed the same Likelihood Ratio Chi Square statistical analysis with respect to the 110 GA sites present in the 1.3-kb region of *pol* to determine whether copackaging of A3G-YFP with A3F-mCH or A3H-mCH influenced the distribution of mutations at GA sites (Figure 3D and E, respectively). We found that the distribution of GA edited sites was not homogenous, indicating that A3G-YFP, A3F-mCH and A3H-mCH induced mutations at some GA target sites at higher frequencies than other GA sites (Figure 3D and E; $^{**}P < 0.0001$). Next, we performed group comparisons between A3G-YFP versus A3F-mCH, A3G-YFP versus A3H-mCH and A3F-mCH versus A3H-mCH, and found that all groups exhibited significant differences (Figure 3D, E and Supplementary Figure S2C, re-

Table 1. Hypermutations induced by APOBEC3 proteins during a single cycle infection in cell culture

APOBEC3	No. of clones sequenced	Hypermutated sequences		GA-to-AA mutations			GG-to-AG mutations		
		Total or dual	Number (%)	Total	Per clone[a]	Per site[b] $\times 10^{-3}$	Total	Per clone[a]	Per site[b] $\times 10^{-3}$
A3G-YFP only	171	Total	164 (96.0)	691	3.9 ± 1.4	38.3	5481	32.1 ± 4.8	428.5
A3F-mCH only	160	Total	149 (93.0)	2348	13.0 ± 6.1	143.3	400	2.2 ± 1.1	34.4
A3D-mCH only	122	Total	11 (9.0)	34	3.4 ± 4.0	28.1	20	1.6 ± 1.2	23.3
A3H-mCH only	122	Total	73 (60.0)	335	4.6 ± 0.4	41.7	53	0.6 ± 0.6	9.3
A3G-YFP + A3F-mCH	250	Total	226 (90.4)	2440	11.1 ± 5.8	98.1	6423	28.3 ± 5.2	364.4
		Dual	112 (49.6)	1954	16.4 ± 4.9	158.6	3426	31.6 ± 4.5	392.2
A3G-YFP + A3D-mCH	122	Total	113 (92.6)	660	5.9 ± 0.9	53.1	3847	34.0 ± 6.6	436.5
		Dual	2 (1.8)	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]
A3G-YFP + A3H-mCH	127	Total	119 (93.7)	1239	10.1 ± 4.6	94.7	3453	28.7 ± 3.6	372.0
		Dual	58 (48.7)	959	17.1 ± 2.2	150.3	1731	33.1 ± 5.9	382.6

[a]The average number of mutations per clone ± standard deviation for three to six independent experiments is shown. For the A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH and A3G-YFP + A3H-mCH groups, the total hypermutated sequences as well as dual-hypermutated sequences were analyzed to determine the average number of mutations per clone.
[b]Mutations/site = total mutations/[sites/clone × total no. of hypermutated sequences]. For the A3G-YFP + A3F-mCH, A3G-YFP + A3D-mCH and A3G-YFP + A3H-mCH groups, Mutations/site were calculated using either total hypermutated sequences or dual-hypermutated sequences and both are shown. Each clone contained 78 GG and 110 GA editing target sites.
[c]Due to the limited number of dual-hypermutated sequences obtained for A3G-YFP + A3D-mCH group, further analysis was not carried out. n.d.; not determined.

Table 2. Effect of nucleotide at +2 position on APOBEC3-induced G-to-A mutation frequencies in cell culture

APOBEC3	Trinucleotide context	Nucleotide at +2 position							
		A		C		G		T	
		Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]	Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]	Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]	Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]
A3G-YFP	GAn	22.2	0.16	1.9	0.01	136.7	1.00	21.2	0.16
	GGn	437.3	0.69	51.8	0.08	629.9	1.00	366.7	0.58
A3F-mCH	GAn	197.8	1.00	41.0	0.21	117.3	0.59	126.9	0.64
	GGn	61.9	1.00	1.1	0.02	14.6	0.24	15.3	0.25
A3D-mCH	GAn	28.5	0.55	9.6	0.18	19.1	0.37	51.9	1.00
	GGn	40.4	1.00	0.0	0.00	11.9	0.29	13.0	0.32
A3H-mCH	GAn	61.0	1.00	13.0	0.21	31.0	0.51	30.7	0.50
	GGn	12.2	0.97	0.0	0.00	12.5	1.00	0.0	0.00
A3G-YFP + A3F-mCH	GAn	196.4	0.90	36.2	0.17	217.6	1.00	124.1	0.57
	GGn	396.8	0.68	46.1	0.08	585.8	1.00	325.3	0.56
A3G-YFP + A3D-mCH	GAn	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]
	GGn	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]	n.d.[c]
A3G-YFP + A3H-mCH	GAn	179.5	0.82	39.0	0.18	218.7	1.00	118.2	0.54
	GGn	398.9	0.75	54.6	0.10	534.5	1.00	362.1	0.68

[a]For single APOBEC3-containing experiments, Mutations/site = total mutations/[sites/clone × total no. of hypermutated clones]; for the combination experiments, Mutations/site = total mutations/[sites/clone × no. of dual-hypermutated clones]. Each clone contained 78 GG and 110 GA editing target sites. The numbers of sites per clone were as follows: GAA, 51; GAC, 19; GAG, 19; GAT, 21; GGA, 36; GGC, 12; GGG, 23; GGT, 7.
[b]Relative mutation frequencies were calculated by setting the highest mutation per site frequency to 1.0, and determining the mutation frequencies for the other nucleotides relative to the highest mutation frequency.
[c]Because only two clones were dual-hypermutated sequences, further analysis was not carried out. n.d., not determined.

spectively; $P < 0.0001$). Thus, the APOBEC3 proteins exhibited unique preferences for GA sites that were targeted for mutation. Similar analysis demonstrated that the distribution of the GA mutations between A3G-YFP versus A3G-YFP + A3F-mCH and A3G-YFP versus A3G-YFP + A3H-mCH were also not homogenous (Figure 3D and E, respectively; $P < 0.0001$). In addition, we also found a significantly different distribution of the GA mutations between A3F-mCH versus A3G-YFP + A3F-mCH and A3H-mCH versus A3G-YFP + A3H-mCH samples (Figure 3D and E, respectively; $^{**}P < 0.0001$). Taken together, these results suggest that each APOBEC3 protein mutates a unique set of GA sites, and when two APOBEC3 proteins are co-packaged, they both contribute to mutations at GA sites and generate distinct patterns of mutations at GA sites.

Local sequence contexts of APOBEC3 protein-mediated mutations of HIV-1 proviral DNA

The target sequences of APOBEC3 protein-mediated mutations are not random and differences in the distribution of mutations as well as local consensus sequences can be discerned (Figure 3D, E and Supplementary Figure S2A and B) (44–46,48,67,74). Biochemical and structural studies have suggested that the amino acids of polynucleotide cytidine deaminases that interact with the substrate single-stranded DNA and the conformational changes in the substrate may play critical roles to permit the reactive cytidine and the immediate neighboring −1, +1 and +2 nucleotides to access the active site (75,76). To systematically investigate the local sequence contexts in which G-to-A mutations occurred, the mutation frequencies in the context of each base found between −5 to +5 positions relative to the

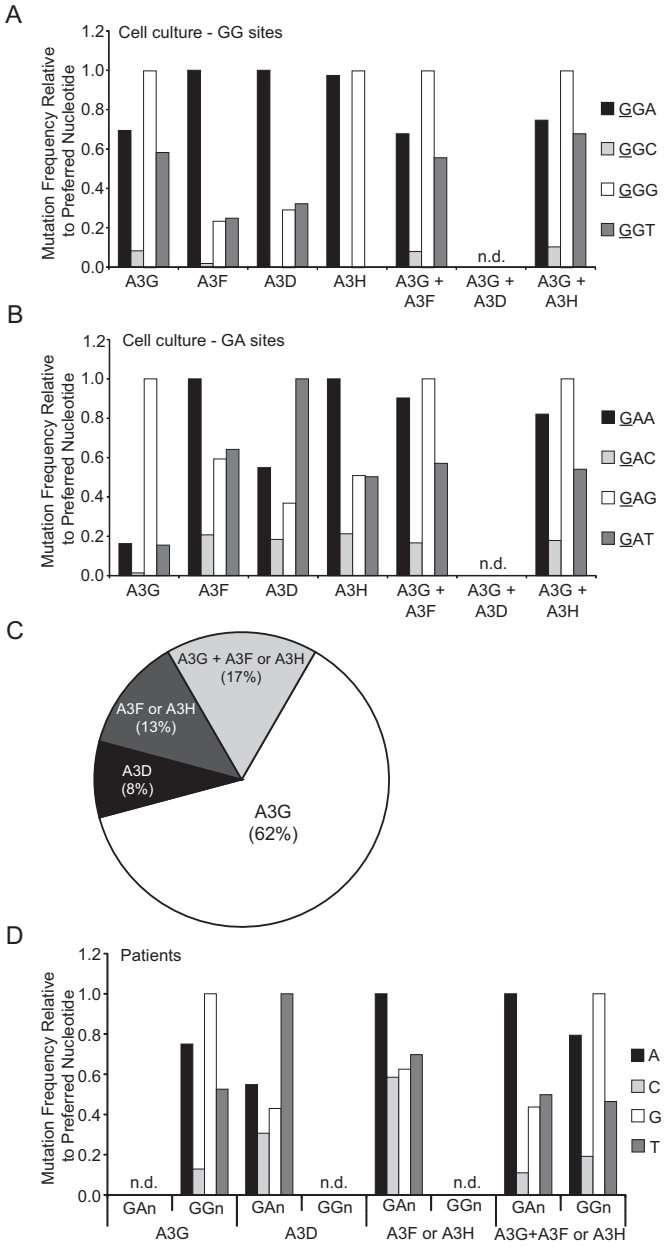

Figure 4. Effect of nucleotide at +2 position on relative mutation frequencies at GA and GG sites in cell culture and in patients. The relative mutation frequencies in cell culture (**A** and **B**) are shown in Table 2. Relative preference of nucleotides at the +2 position in both the GG (A) and GA (B) edited sites were plotted for A3G-YFP only, A3F-mCH only, A3D-mCH only, A3H-mCH only, A3G-YFP + A3F-mCH and A3G-YFP + A3H-mCH conditions in the cell culture experiments. For the combination experiments only the dual-hypermutated sequences were analyzed and for A3G-YFP + A3D-mCH combination experiments, only two clones were identified as dual hypermutated and hence the analysis was not determined (n.d.). For all conditions, the mutation frequency of the most preferred nucleotide was set to 1.0 and the mutation frequencies for the other nucleotides relative to the most preferred nucleotide were shown. (**C**) The contribution of the different APOBEC3 proteins in patients as determined by the proportion of hypermutated sequences. (**D**) Similar analysis shown in panel A and B was performed to determine relative mutation frequencies for hypermutated sequences isolated from patients and the sequences fall into four categories that fulfill the criteria for A3G only, A3D only, A3F or A3H only, and A3G + A3F (or A3H) induced G-to-A mutations. All the comutated sequences identified in patients have local sequence context preference of A3G + (A3F or A3H) pattern; none for A3G + A3D were identified. The relative mutation frequencies at GA edited sites for sequences that primarily had GG site mutations (A3G pattern) and at GG edited sites for sequences that primarily had GA site mutations (A3D or A3F/A3H) were not determined (n.d.) because of the limited numbers. Data were analyzed with generalized linear model and loglinear modeling techniques. The significantly preferred nucleotide in the +2 positions are indicated (*$P < 0.05$).

deaminated cytidine were interrogated. We found distinct nucleotide preference hierarchy at the +2 position for each APOBEC3 protein relative to the guanine nucleotide that was mutated to adenine (underlined) in 5′-GA and 5′-GG editing sites upon deamination of the cytidine in the minus-strand DNA (Table 2). We first determined the frequencies of mutations per site for each nucleotide (A, G, C or T) at the +2 position (mutations/[sites per sequence × number of hypermutated sequences analyzed]). We then ranked the mutation frequencies and set the highest mutation frequency to 1.0, which was then used to calculate the relative frequencies of mutations for the other nucleotides. Using the numbers of G-to-A mutations in the trinucleotide context when the +2 nucleotide is an A, C, G or T, we carried out Fisher's Exact tests (with Bonferroni correction for multiple comparisons) to determine whether the highest mutation frequency in a specific +2 nucleotide context was significantly different from the other three +2 nucleotide contexts. The results obtained from cell culture experiments are shown in Table 2, Figure 4A and B. The mutation frequencies induced by A3G in the GG and GA contexts are highest and significantly different ($P < 0.05$) when the +2 nucleotide is G (Figure 4A and B, respectively); the mutation frequencies are significantly lower when the +2 nucleotide is A or T (Figure 4A and Figure B, respectively). Overall A3G is less likely to mutate GG or GA sites when C is at the +2 nucleotide position ($P < 0.05$) but A3G can accommodate A or T at the +2 position much more frequently in the GG context compared to the GA context.

The most favored nucleotide at +2 position by A3F and A3H for inducing mutations at the GA sites is A ($P < 0.05$), followed by G or T (Figure 4B). Unlike A3F and A3H, A3D prefers a T at the +2 position followed by A or G; however, the differences are not statistically significant because very few A3D-induced GA mutations were observed. Interestingly, all APOBEC3 cytidine deaminases highly disfavored GA editing sites when a C was at the +2 position (Figure 4B).

Next, we compared the relative nucleotide preferences at the +2 position of G-to-A mutations when A3G-YFP + A3F-mCH or A3G-YFP + A3H-mCH were copackaged into the same virus particles (Table 2; Figure 4A and B). At the GG edited sites, the preferences for nucleotides at the +2 position were nearly identical to the preferences observed when A3G alone was present in the virus particles, further confirming that mutations in the GG context are primarily induced by A3G irrespective of the presence or absence of other APOBEC3 proteins (Table 2 and Figure 4A). Interestingly, the preference for nucleotides at the +2 position in the GA edited sites in the dual-hypermutated proviruses (A3G-YFP + A3F-mCH or A3G-YFP + A3H-mCH) were not

similar to the A3G-YFP, A3F-mCH or A3H-mCH alone (Table 2 and Figure 4B; $P < 0.0001$ for all group comparisons between A3G-YFP versus A3G-YFP + A3F-mCH; A3G-YFP versus A3G-YFP + A3H-mCH, A3F-mCH versus A3G-YFP + A3F-mCH; A3H-mCH versus A3G-YFP + A3H-mCH), indicating that both copackaged APOBEC3 proteins contributed to the observed mutations at the GA sites.

Finally, using the +2 nucleotide preference, we investigated whether the synergistic increase in GA mutations observed when A3G-YFP and A3H-mCH were copackaged was the result of an increase in the deaminase activity of A3H and/or A3G at GA sites. A3G and A3H prefer GAG and GAA sites for inducing mutations at GA sites, respectively (Figure 4B and Supplementary Figure S3). When both A3G and A3H were co-expressed, we observed a synergistic increase in GA mutations at the GAA sites that are preferred by A3H ($P = 0.027$), but not the GAG sites that are preferred by A3G ($P > 0.05$), suggesting that the increase in GA mutations was due to a synergistic increase in the deaminase activity of A3H but not A3G. In contrast, we observed only an additive increase in GA mutations when A3G and A3F were copackaged; as expected, we observed additive activity at the GAG (A3G preferred sites) and GAA (A3F preferred sites) contexts when A3G-YFP and A3F-mCH were copackaged (Supplementary Figure S3; *P > 0.05*).

Frequent comutation of proviruses by APOBEC3 proteins in HIV-1 infected patients

In PBMCs isolated from HIV-1 infected patients, proviral DNAs hypermutated by APOBEC3 proteins have been observed. The existence of GG and GA mutations in the same genome has also been described; however, none of the studies have shown that they were the result of copackaged APOBEC3 proteins (50,77–79). In addition, one study analyzed HIV sequences from the Los Alamos database (22), and suggested that A3G and A3F rarely comutated the same genome *in vivo*. Thus, we sought to determine whether APOBEC3 proteins are copackaged into the same HIV-1 virions and comutate the same viral genome in patients. To address this question, we analyzed APOBEC3-mediated hypermutated sequences obtained from a well-characterized cohort of seven chronically infected HIV-1 patients. Clinical characteristics of the seven patients are shown in Supplementary Table S1. The estimated duration of infection ranged from 1.2 to 21.8 years and all of the study participants had detectable viremia with a median viral load of 4.8 $\log_{10}$ copies/ml at the time of sampling (Supplementary Table S1). A genetic screen for A3H haplotype was determined as described before (6,68) and the poorly expressed hap I was found in one patient (patient 4), hap II, the stably expressed and potent antiviral variant of A3H, was found in six patients of whom one (patient 10) was homozygous at this locus and the remaining five were heterozygous (Supplementary Table S2).

We used the local sequence contexts determined in the cell culture experiments described above to determine whether similar sequence contexts were also observed *in vivo*. We performed SGS of the 1.1-kb fragment of the *pro-pol* region of HIV-1 plasma viral RNA as well as proviral DNA extracted from PBMCs (Table 3). We and others have observed that as a result of purifying selection, there are few if any hypermutated sequences in viral RNA (47,50,80). We obtained a total 217 independent plasma viral RNA sequences, and as expected, the plasma viral RNA sequences had very few G-to-A mutations per clone (GG-to-AG mutations per clone ranged from 0–1 and GA-to-AA mutations per clone ranges from 1–3) (Table 3). Then we generated a non-hypermutated consensus sequence for each patient from these plasma viral RNA sequences to which we aligned all the proviral DNA sequences. We determined a threshold level of GG-to-AG and GA-to-AA mutations that significantly exceeded levels of these mutations detected in the corresponding plasma viral RNA (Fisher's Exact tests for each patient; $P < 0.05$); for each patient, threshold levels for hypermutation ranged from 5 to 7 for GG-to-AG and 7 to 10 GA-to-AA mutations compared to the consensus plasma viral RNA sequence for each patient.

We then characterized the nature of G-to-A hypermutations in patients by analyzing a total of 328 proviral DNA sequences (17–113 sequences per patient), of which 24 (7.3%) were hypermutated (Tables 3 and 4). This frequency of hypermutation is comparable to the minimal estimate of hypermutation reported previously (47,50). Only one of the seven patients did not have any hypermutated sequences even though 52 proviral DNA sequences were obtained (Table 3). Hypermutants that predominantly had mutations at GG sites but also had a high frequency of GA site mutations (>10% of GG site mutations) were defined as dual-hypermutated proviruses. Similarly, hypermutants that predominantly had GA site mutations but also had a high frequency of GG site mutations (17% of GA site mutations) were defined as dual-hypermutated proviruses. Using these criteria, 62% of the hypermutated proviral sequences were mutated by A3G only (15/24) and 21% of the hypermutated proviruses were mutated by A3F, A3H or A3D). The remaining 17% (4/24) of proviruses were dual-hypermutated, indicating that copackaging of APOBEC3 proteins and comutation of the same viral genome can occur *in vivo* (Table 4 and Figure 4C).

In addition to A3F, A3D and certain haplotypes of A3H can predominantly induce mutations in the GA context and potentially contribute to the overall GA mutations in patients (7,68,81). In the cell culture experiments described above, we observed that A3F and A3H exhibit a similar preference for nucleotides at the +2 position, whereas A3D exhibits a distinct preference; we analyzed the patient-derived hypermutated sequences to determine which APOBEC3 protein contributed to the GA site mutations. The results indicated that 8% of the hypermutated proviruses were mutated by A3D only and 13% were mutated by A3F and/or A3H (Table 4; Figure 4C and D). One patient only expressed the unstable A3H Hap I, and its A3H presumably could not contribute to the GA site mutations; however, the hypermutated proviruses derived from this patient were predominantly mutated by A3G, which excluded the possibility of gaining further insights into the identity of the APOBEC3 proteins that contributed to GA site mutations (Table 3 and Supplementary Table S2). Importantly, our results suggest that A3G and A3F (or A3H) were co-

Table 3. Single-genome sequencing analysis of viral RNA and cellular DNA from patients

	Viral RNA		Cellular DNA	
Patient	Total sequences	Hypermutated sequences[a]	Total sequences	Hypermutated sequences[a]
1	24	0	52	0
2	50	0	17	5
4	27	0	22	3
7	16	0	44	8
8	21	0	21	2
10	71	0	59	2
11	8	0	113	4
Total	217	0	328	24

[a]A consensus plasma viral RNA sequence was generated for each patient. The average number of G-to-A changes between individual RNA sequences and the consensus RNA sequence was used to determine the number of G-to-A changes needed to achieve statistical significance. A sequence that had a significantly higher number of mutations than the plasma viral RNA consensus sequence was considered a hypermutant (Fisher's Exact test; $P < 0.05$).

Table 4. Hypermutations induced by APOBEC3 cytidine deaminases in patients

		GA-to-AA mutations			GG-to-AG mutations		
APOBEC3	No. of Seq.	Total	Per sequence[a]	Per site ($\times 10^{-3}$)[b]	Total	Per sequence[a]	Per site ($\times 10^{-3}$)[b]
A3G	15	81	5.4 ± 3.5	51.8	454	30.3 ± 10.2	457.2
A3D	2	55	27.5 ± 27.6	241.2	5	2.5 ± 3.5	40.3
A3F or A3H	3	29	9.7 ± 2.1	86.8	3	1.0 ± 1.0	15.5
A3G + A3F or A3H	4	57	14.3 ± 4.6	132.3	131	32.8 ± 16.5	490.6

[a]The average number of mutations per sequence ± standard deviation was calculated for each group.
[b]Mutations/site = total mutations/[sites/sequence × no. of sequences]. The numbers of sites per sequence were determined using the plasma viral RNA consensus sequence for each patient.

packaged into the same virions and comutated the same viral genomes.

Next, we compared the effects of nucleotides at the +2 position on the frequencies of hypermutation in patients (Table 5 and Figure 4D). The frequencies of mutations per site for each nucleotide context were determined, the highest frequency was set to 1.0, and the relative mutation frequencies for the other nucleotides were determined as described above for the cell culture experiments (Figure 4A and B). The patterns of relative mutation frequencies induced by A3G in the GG context in patients and in cell culture were similar but not identical ($P < 0.041$); a G was the preferred nucleotide at the +2 position both in cell culture and patients (compare Figure 4A and D; $P < 0.0005$). The concordance between the patient studies and cell culture results confirmed that most of the GG mutations in patients were primarily induced by A3G cytidine deaminase activity. Moreover, the overall local sequence context preference at GA edited sites by A3D and A3F/A3H deaminases in patients displayed very similar patterns compared to the patterns observed in cell culture when they were expressed alone (compare Figure 4B and D; $P > 0.05$). In both patients and in cell culture, A3F and A3H preferred an adenine and A3D preferred a thymidine at the +2 position in the GA edited site context ($P > 0.05$), and cytidine was disfavored at the +2 position in patients as well as in cell culture ($P < 0.05$).

Next, we analyzed the nucleotide preference at the +2 position in the dual-hypermutated sequences derived from patients (Table 5 and Figure 4D). The preferred nucleotides at the +2 position in the GG context when A3G and A3F (or A3H) deaminases were copackaged into virions in cell culture experiments and patients were similar (compare Figure 4A and D, respectively; $P > 0.05$), confirming that the GG site mutations in dual-hypermutated sequences derived from patients are primarily induced by A3G. Furthermore, the preferred nucleotides at the +2 position in the GA edited sites of dual-hypermutated patient-derived sequences were not significantly different from that observed in cell culture experiments performed with A3G-YFP + A3F-mCH or A3G-YFP + A3H-mCH (compare Figure 4B and D; $P > 0.05$). These results are consistent with the interpretation that most of the dual-hypermutated sequences were derived from the deamination activity of A3G as well as A3F or A3H.

DISCUSSION

In this study, we demonstrate for the first time that different APOBEC3 proteins can efficiently copackage and comutate the same viral genomes in a single cycle of HIV-1 replication. Moreover, we showed that comutation can be observed in proviral DNA recovered from PBMCs of chronically infected HIV-1 patients, suggesting that A3G and other APOBEC3 proteins can copackage and comutate HIV-1 *in vivo*. Using single-virion fluorescence microscopy, we observed that A3D, A3F and A3H can be copackaged with A3G into the same virions; furthermore, the amounts of the APOBEC3 proteins packaged into virions was similar to the amounts in virions containing a single APOBEC3 variant, suggesting that the APOBEC3 proteins do not compete for virion encapsidation. Previous studies by us and others concluded that APOBEC3 proteins are packaged into virions by interacting with viral and nonviral RNAs (25,28–32), and the amounts of APOBEC3 pro-

Table 5. Effect of nucleotide at +2 position on mutation frequencies *in vivo*

		Nucleotide at +2 position							
		A		C		G		T	
APOBEC3 proteins	Dinuc. Seq.	Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]	Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]	Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]	Mut./site[a] ($\times 10^{-3}$)	Rel. Mut. Freq.[b]
A3G	GAn	52.8	n.d.[c]	10.7	n.d.[c]	103.7	n.d.[c]	43.4	n.d.[c]
	GGn	471.5	0.75	80.8	0.13	628.8	1.00	330.6	0.53
A3D	GAn	223.4	0.55	125	0.31	175	0.43	407.4	1.00
	GGn	44.1	n.d.[c]	0.0	n.d.[c]	0.0	n.d.[c]	111.1	n.d.[c]
A3F or A3H	GAn	110.3	1.00	64.5	0.58	69.0	0.63	76.9	0.70
	GGn	9.5	n.d.[c]	58.8	n.d.[c]	21.3	n.d.[c]	0.0	n.d.[c]
A3G + A3F or A3H	GAn	216.4	1.00	23.8	0.11	94.6	0.44	107.8	0.50
	GGn	517.7	0.79	125.0	0.19	652.2	1.00	303.0	0.46

[a]Mutations/site = total mutations/[sites/sequence × no. of sequences]. The numbers of sites per sequence were determined using the plasma viral RNA consensus sequence for each patient.
[b]The relative mutation frequencies were calculated by setting the highest mutations/site frequency to 1 and determining the mutations/site frequency for the remaining nucleotides relative to the highest frequency.
[c]The hypermutated sequences exhibited other nucleotide variation relative to the patient consensus sequences. Although we observed GA mutations in sequences that predominantly had GG mutations (A3G Only), the frequencies of these mutations were not significantly higher than those observed for non-hypermutated sequences. Therefore, we were unable to determine whether the GA mutations observed in sequences predominantly hypermutated by A3G, and GG mutations observed in sequences predominantly hypermutated by A3F (or A3D or A3H) were due to A3G- or A3F-(or A3D- or A3H-) induced deamination, respectively, or whether these mutations arose during multiple rounds of viral replication in the patients. Consequently, we did not analyze the relative mutation frequencies of these mutations. n.d., not determined.

teins packaged into virions is proportional to the amount of packaged RNA. Furthermore, a recent analysis using cross-linking and immunoprecipitation technique concluded that the promiscuity and high affinity of these proteins to bind to both viral and non-viral RNAs ensures their virion incorporation (36). Although promiscuous binding to RNA provides a potential mechanism for independent packaging of APOBEC3 proteins, the APOBEC3 proteins could compete for virion incorporation at a step prior to their interaction with RNA. The results of our studies directly determined that APOBEC3 proteins are copackaged and do not compete for packaging into virions.

Previous studies have observed GG and GA mutations in the same viral genomes (50,69,77–79), which can be explained by three potential mechanisms. First, copackaging and comutation of the viral genomes may have occurred in a single cycle of replication; second, comutation could be the result of hypermutation by different APOBEC3 proteins in sequential rounds of replication; and third, one APOBEC3 protein could induce mutations in both GG and GA contexts. The previous studies could not distinguish between these possibilities, in part because the potential for APOBEC3 proteins to induce mutations at both GG and GA sites was not fully appreciated. As the results of these and other studies show, A3G can induce mutations at GA sites at ∼12% the frequency with which it induces mutations at GG sites (7,48,67). We also observed that A3F and A3H can induce mutations GG sites at 17 and 13% the frequency with which they induce mutations at GA sites, respectively. Thus, it is necessary to determine the ratio of mutations at GG and GA sites and determine whether the frequency of mutations at GA sites is higher than expected by individual APOBEC3 proteins in order to conclude that another APOBEC3 contributed to the hypermutation of genomes containing mutations at both GG and GA sites.

Another study analyzed HIV-1 sequences from the Los Alamos HIV database and concluded that comutation of the same HIV-1 genomes by A3G and A3F occurs rarely (22). In this study, a bioinformatics-based determination of G-to-A hypermutations within the same sequence using only the dinucleotide GG and GA motifs representation approach was employed to analyze different groups and subtypes of HIV-1. This discrepancy could be due to the high genetic variation in the HIV-1 population, which limits the utility of population-based sequence analysis; the lack of a consensus sequence for each patient to align patient-derived reads may have reduced the sensitivity of identification of comutated genomes. In our study, we analyzed single-genome sequences and generated a consensus sequence for each patient based on the viral RNA sequences, potentially providing a more sensitive assay for hypermutation. We analyzed single-genome proviral DNA sequences from seven viremic chronically infected HIV-1 patients and mapped these APOBEC3-mediated G-to-A hypermutations onto individual proviral DNA reads by aligning to a consensus sequence generated for each patient based on the non-hypermutant plasma viral RNA sequences.

In these studies, we systematically analyzed the +1 and +2 nucleotide positions of G-to-A mutations induced by various APOBEC3 proteins and found distinct trinucleotide contexts for each APOBEC3 protein. Importantly, we used these trinucleotide sequence contexts to identify the APOBEC3 proteins that induced hypermutation in the proviral DNAs derived from patients. The trinucleotide contexts we observed for A3G and A3F are in agreement with recently described trinucleotide contexts (48). In contrast to the previous study, we performed more detailed analysis of the relative frequencies of G-to-A mutations in various trinucleotide contexts for each APOBEC3 protein. In addition, in contrast to the previously described preference of A3H for only GGA sites, we observed a strong preference for both GGA and GGG contexts. Our sequence contexts for A3D also differed in that we observed a preference for GAT sites, while the previous study did not provide any contexts for GA sites but indicated a preference for GGA sites. The reasons for these differences are not clear, but may be due to the low deaminase activity of A3D, which may have resulted in very few G-to-A mutations above the background in the previous study.

Combinations of A3G-YFP + A3F-mCH or A3G-YFP + A3H-mCH resulted in either no increase or in a 1.8-

fold increase, respectively, in the frequency of mutations at GA sites compared to the additive increase expected from the deaminase activity of both APOBEC3 proteins. These results suggest that A3G-YFP + A3H-mCH, but not A3G-YFP + A3F-mCH, act synergistically to increase mutations at GA sites. A direct or indirect interaction between A3G and A3H could potentially result in synergism and an increase in GA mutations. A direct interaction could involve heterodimerization between the two deaminases and enhancement of the activity of one or both enzymes at 5′-TC sites. Alternatively, an indirect interaction, such as inhibition of the rate of plus-strand DNA synthesis (46,70,72,82,83), could allow the deaminases to act on the substrate minus-strand DNA for a longer period of time, leading to an increase in the frequency of GA mutations. Since we observed a synergistic increase in the activity of A3H but not A3G, the results suggest that the mechanism does not involve inhibition of reverse transcription, which would be expected to increase the deamination activity of both enzymes. However, additional studies are needed to understand the underlying molecular mechanism(s) that contributes to the synergistic increase in A3H deamination activity when copackaged with A3G. In our cell culture based assays, we observed an additive increase, but not a synergistic increase, in the overall antiviral activity when A3G and A3H were co-expressed in virus producing cells. While increased GA mutations would lead to more lethal mutations, we and others have recently concluded that hypermutation usually results in lethal mutagenesis (44,84), while others have concluded that hypermutation can result in sublethal mutagenesis and contribute to viral genetic variation (48). It remains possible that a twofold synergistic increase in the activity of A3H may lead to a small increase in antiviral activity that was not apparent in our single-cycle infectivity assays. Another potential outcome of copackaging and comutation is that the resulting proviruses may possess a greater number of G-to-A mutations, and therefore more stop codons, which would increase their potential for expression of aberrant protein products that can potentially be recognized by cytotoxic T lymphocytes (CTLs), facilitating CTL-mediated clearance of infected cells (85,86). It has been reported that hypermutation by APOBEC3 is associated with strong CTL and natural killer cell responses (87,88); additional studies are needed to determine whether copackaging and comutation can potentially facilitate CTL responses that eliminate infected cells (48,89).

The observed synergism between copackaged A3G and A3H may provide insights into the mechanisms by which the APOBEC3 proteins interact with each other, move processively on the substrate DNA and carry out cytidine deamination. It is also possible that A3D, A3F and A3H can be copackaged, and act synergistically to increase GA mutations in a manner that leads to a synergistic increase in their antiviral activity. However, A3F and A3H both prefer GAA sites and it is not possible to clearly define which APOBEC3 protein or proteins contributed to the GA mutations in proviruses that are predominantly mutated at GA sites. While A3D is unique in its preference for GAT sites, A3F also induces mutations at GAT sites at a high frequency, and at present, it is not possible to clearly identify hypermutated genomes that were dual-hypermutated by A3F/A3H and A3D. Additional studies are needed to further define the sequence contexts that are preferred by A3D, A3F and A3H proteins and to determine whether they can be copackaged and can comutate the same genomes.

Our cell culture-based data clearly demonstrate that ∼50% of the hypermutated sequences retrieved from cells infected with viruses produced in the presence of A3G-YFP + A3F-mCH or A3G-YFP + A3H-mCH were dual-hypermutated, which is consistent with the copackaging efficiency. Analysis of proviral genomes isolated from PBMCs of chronically infected HIV-1 patients indicated that some genomes were hypermutated at both GG and GA sites, suggesting that A3G and A3F (or A3H/A3D) were copackaged into the same virions and induced comutation. One potential explanation for the co-existence of GG and GA mutations in the same genome in HIV-1 patients is the accumulation of G-to-A mutations induced by different APOBEC3 proteins during different cycles of replication *in vivo*. However, comutation at GG and GA sites in sequential rounds of viral replication is unlikely for the following reasons. First, our recent studies indicate that hypermutation by one APOBEC3 protein is likely to lead to lethal hypermutation of the viral genome (84). Results from previous studies (47,50,80) as well as from the current study showed that because of purifying selection very few hypermutated genomic RNAs are packaged into virions, making it unlikely that RNAs from hypermutated genomes would be copackaged with another APOBEC3 protein during subsequent rounds of infections. Second, packaging and replication of the hypermutated genomes would require superinfection of the hypermutated provirus-containing cells with another replication-competent virus, which has been shown to be a rare event (90). Third, hypermutation itself is an infrequent event; for example, in our study only 24 of 328 (7.3%) proviral genomes from PBMCs of seven viremic chronically-infected HIV-1 patients were hypermutated, consistent with the previously reported hypermutation frequency (47,50). Since sequential rounds of hypermutation would require all three of these rare events, we believe that proviruses with both GG and GA mutations most likely resulted from copackaging of multiple APOBEC3 proteins and comutation of the same viral genomes in a single round of replication. Additionally, since sequential rounds of hypermutation are likely to be rare, we would expect the frequency of comutation to be less than the frequency of hypermutation by A3F or A3H alone. However, we observed a similar frequency of A3G and A3F/A3H-mediated comutation (4 of 24 sequences; ∼17%) and A3F (or A3H) only induced hypermutation (3 of 24 sequences; ∼13%), which suggests that comutation predominantly occurred through copackaging of different APOBEC3 proteins in a single round of replication.

Specific sequences and/or substrate nucleic acid secondary structure could potentially affect target site specificities of APOBEC3 proteins. We analyzed 1320 nucleotides of the HIV-1 RT region, which represents ∼14% of the viral genome. Since APOBEC3 proteins are likely to interact with very short regions near the target nucleotide, we believe analysis of a large portion of the genome provides an average of the sequences and RNA structures that are likely to be encountered throughout the HIV-1 genome and in dif-

ferent HIV-1 isolates. Our experiments were performed with HDV-eGFP, which does not express HIV-1 accessory proteins Vif, Vpu, Vpr and Nef. Therefore, it is possible that the accessory proteins, Vpr, Vpu and Nef could influence the target DNA specificity of APOBEC3 proteins, and the target DNA specificities observed in infected individuals might be different from those determined in our cell-based assays. However, to our knowledge there is no direct evidence in the literature suggesting that these accessory proteins influence the deaminase activities and/or target site specificities of APOBEC3 proteins. In addition, we observed that the target site specificity of A3G in infected individuals and in cell-based assays was very similar, suggesting that the presence of the other accessory proteins did not influence A3G's target site specificity.

In summary, the results presented here show that different APOBEC3 proteins can copackage into the same virions independently of each other and comutate the same viral genomes. Analysis of chronically infected patients led to the identification of dual-hypermutated proviruses, providing strong evidence that copackaging of these deaminases and comutation of HIV-1 genomes can frequently occur *in vivo*. The data reported here, which demonstrate that A3G and A3F or A3H proteins can inactivate the same genomes, indicate that they all are co-expressed co-ordinately in sufficient amounts in individual HIV target cells and copackaged to inhibit HIV replication. Furthermore, the findings suggest the possibility for cooperative inhibition of invading viral pathogens by multiple APOBEC3 proteins *in vivo*. Additional studies are needed to determine how the immune activation status (91–93), which may impact IFN induction (20,21,91,92,94) and APOBEC3 protein expression (20,93–95), affects copackaging and comutation of HIV-1 genomes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank members of the laboratories of Vinay K. Pathak and Wei-Shau Hu for valuable discussion of results. We are grateful to the donor participants for their generous contributions to this study.

FUNDING

This work was supported in part by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research, and by Intramural AIDS Targeted Antiviral Program grant funding to V.K.P. Funding for open access charge: National Cancer Institute; National Institutes of Health, USA.
Conflict of interest statement. None declared.

REFERENCES

1. Zheng,Y.H., Jeang,K.T. and Tokunaga,K. (2012) Host restriction factors in retroviral infection: promises in virus-host interaction. *Retrovirology*, **9**, 112.
2. Malim,M.H. and Bieniasz,P.D. (2012) HIV restriction factors and mechanisms of evasion. *Cold Spring Harb. Perspect. Med.*, **2**, a006940.
3. Desimmie,B.A., Delviks-Frankenberrry,K.A., Burdick,R.C., Qi,D., Izumi,T. and Pathak,V.K. (2014) Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. *J. Mol. Biol.*, **426**, 1220–1245.
4. Simon,V., Bloch,N. and Landau,N.R. (2015) Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. *Nat. Immunol.*, **16**, 546–553.
5. Sheehy,A.M., Gaddis,N.C., Choi,J.D. and Malim,M.H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, **418**, 646–650.
6. OhAinle,M., Kerns,J.A., Malik,H.S. and Emerman,M. (2006) Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H. *J. Virol.*, **80**, 3853–3862.
7. Hultquist,J.F., Lengyel,J.A., Refsland,E.W., LaRue,R.S., Lackey,L., Brown,W.L. and Harris,R.S. (2011) Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J. Virol.*, **85**, 11220–11234.
8. Harari,A., Ooms,M., Mulder,L.C. and Simon,V. (2009) Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H. *J. Virol.*, **83**, 295–303.
9. Wiegand,H.L., Doehle,B.P., Bogerd,H.P. and Cullen,B.R. (2004) A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J.*, **23**, 2451–2458.
10. Mangeat,B., Turelli,P., Caron,G., Friedli,M., Perrin,L. and Trono,D. (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*, **424**, 99–103.
11. Sheehy,A.M., Gaddis,N.C. and Malim,M.H. (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat. Med.*, **9**, 1404–1407.
12. Yu,X., Yu,Y., Liu,B., Luo,K., Kong,W., Mao,P. and Yu,X.F. (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science*, **302**, 1056–1060.
13. Marin,M., Golem,S., Rose,K.M., Kozak,S.L. and Kabat,D. (2008) Human immunodeficiency virus type 1 Vif functionally interacts with diverse APOBEC3 cytidine deaminases and moves with them between cytoplasmic sites of mRNA metabolism. *J. Virol.*, **82**, 987–998.
14. Dang,Y., Wang,X., Esselman,W.J. and Zheng,Y.H. (2006) Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *J. Virol.*, **80**, 10522–10533.
15. Yu,Q., Chen,D., Konig,R., Mariani,R., Unutmaz,D. and Landau,N.R. (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J. Biol. Chem.*, **279**, 53379–53386.
16. Refsland,E.W., Hultquist,J.F. and Harris,R.S. (2012) Endogenous origins of HIV-1 G-to-A hypermutation and restriction in the nonpermissive T cell line CEM2n. *PLoS Pathog.*, **8**, e1002800.
17. Ulenga,N.K., Sarr,A.D., Thakore-Meloni,S., Sankale,J.L., Eisen,G. and Kanki,P.J. (2008) Relationship between human immunodeficiency type 1 infection and expression of human APOBEC3G and APOBEC3F. *J. Infect. Dis.*, **198**, 486–492.
18. Koning,F.A., Newman,E.N., Kim,E.Y., Kunstman,K.J., Wolinsky,S.M. and Malim,M.H. (2009) Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J. Virol.*, **83**, 9474–9485.
19. Ying,S., Zhang,X., Sarkis,P.T., Xu,R. and Yu,X. (2007) Cell-specific regulation of APOBEC3F by interferons. *Acta Biochim. Biophys. Sin.*, **39**, 297–304.
20. Pillai,S.K., Abdel-Mohsen,M., Guatelli,J., Skasko,M., Monto,A., Fujimoto,K., Yukl,S., Greene,W.C., Kovari,H., Rauch,A. *et al.* (2012) Role of retroviral restriction factors in the interferon-alpha-mediated suppression of HIV-1 in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 3035–3040.
21. Chen,K., Huang,J., Zhang,C., Huang,S., Nunnari,G., Wang,F.X., Tong,X., Gao,L., Nikisher,K. and Zhang,H. (2006) Alpha interferon potently enhances the anti-human immunodeficiency virus type 1 activity of APOBEC3G in resting primary CD4 T cells. *J. Virol.*, **80**, 7645–7657.
22. Ebrahimi,D., Anwar,F. and Davenport,M.P. (2012) APOBEC3G and APOBEC3F rarely co-mutate the same HIV genome. *Retrovirology*, **9**, 113.
23. Lassen,K.G., Wissing,S., Lobritz,M.A., Santiago,M. and Greene,W.C. (2010) Identification of two APOBEC3F splice variants

displaying HIV-1 antiviral activity and contrasting sensitivity to Vif. *J. Biol. Chem.*, **285**, 29326–29335.
24. Yang,Y., Guo,F., Cen,S. and Kleiman,L. (2007) Inhibition of initiation of reverse transcription in HIV-1 by human APOBEC3F. *Virology*, **365**, 92–100.
25. Wang,T., Tian,C., Zhang,W., Sarkis,P.T. and Yu,X.F. (2008) Interaction with 7SL RNA but not with HIV-1 genomic RNA or P bodies is required for APOBEC3F virion packaging. *J. Mol. Biol.*, **375**, 1098–1112.
26. Zhang,W., Du,J., Yu,K., Wang,T., Yong,X. and Yu,X.F. (2010) Association of potent human antiviral cytidine deaminases with 7SL RNA and viral RNP in HIV-1 virions. *J. Virol.*, **84**, 12903–12913.
27. Bach,D., Peddi,S., Mangeat,B., Lakkaraju,A., Strub,K. and Trono,D. (2008) Characterization of APOBEC3G binding to 7SL RNA. *Retrovirology*, **5**, 54.
28. Svarovskaia,E.S., Xu,H., Mbisa,J.L., Barr,R., Gorelick,R.J., Ono,A., Freed,E.O., Hu,W.S. and Pathak,V.K. (2004) Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. *J. Biol. Chem.*, **279**, 35822–35828.
29. Zennou,V., Perez-Caballero,D., Gottlinger,H. and Bieniasz,P.D. (2004) APOBEC3G incorporation into human immunodeficiency virus type 1 particles. *J. Virol.*, **78**, 12058–12061.
30. Luo,K., Liu,B., Xiao,Z., Yu,Y., Yu,X., Gorelick,R. and Yu,X.F. (2004) Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging. *J. Virol.*, **78**, 11841–11852.
31. Khan,M.A., Kao,S., Miyagi,E., Takeuchi,H., Goila-Gaur,R., Opi,S., Gipson,C.L., Parslow,T.G., Ly,H. and Strebel,K. (2005) Viral RNA is required for the association of APOBEC3G with human immunodeficiency virus type 1 nucleoprotein complexes. *J. Virol.*, **79**, 5870–5874.
32. Wang,T., Tian,C., Zhang,W., Luo,K., Sarkis,P.T., Yu,L., Liu,B., Yu,Y. and Yu,X.F. (2007) 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. *J. Virol.*, **81**, 13112–13124.
33. Alce,T.M. and Popik,W. (2004) APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein. *J. Biol. Chem.*, **279**, 34083–34086.
34. Schafer,A., Bogerd,H.P. and Cullen,B.R. (2004) Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. *Virology*, **328**, 163–168.
35. Khan,M.A., Goila-Gaur,R., Kao,S., Miyagi,E., Walker,R.C. Jr and Strebel,K. (2009) Encapsidation of APOBEC3G into HIV-1 virions involves lipid raft association and does not correlate with APOBEC3G oligomerization. *Retrovirology*, **6**, 99.
36. Apolonia,L., Schulz,R., Curk,T., Rocha,P., Swanson,C.M., Schaller,T., Ule,J. and Malim,M.H. (2015) Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. *PLoS Pathog.*, **11**, e1004609.
37. Izumi,T., Burdick,R., Shigemi,M., Plisov,S., Hu,W.S. and Pathak,V.K. (2013) Mov10 and APOBEC3G localization to processing bodies is not required for virion incorporation and antiviral activity. *J. Virol.*, **87**, 11047–11062.
38. Navarro,F., Bollman,B., Chen,H., Konig,R., Yu,Q., Chiles,K. and Landau,N.R. (2005) Complementary function of the two catalytic domains of APOBEC3G. *Virology*, **333**, 374–386.
39. Iwatani,Y., Takeuchi,H., Strebel,K. and Levin,J.G. (2006) Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. *J. Virol.*, **80**, 5992–6002.
40. Chiu,Y.L., Soros,V.B., Kreisberg,J.F., Stopak,K., Yonemoto,W. and Greene,W.C. (2005) Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells. *Nature*, **435**, 108–114.
41. Chiu,Y.L., Witkowska,H.E., Hall,S.C., Santiago,M., Soros,V.B., Esnault,C., Heidmann,T. and Greene,W.C. (2006) High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 15588–15593.
42. Chaipan,C., Smith,J.L., Hu,W.S. and Pathak,V.K. (2013) APOBEC3G restricts HIV-1 to a greater extent than APOBEC3F and APOBEC3DE in human primary CD4+ T cells and macrophages. *J. Virol.*, **87**, 444–453.
43. Refsland,E.W., Stenglein,M.D., Shindo,K., Albin,J.S., Brown,W.L. and Harris,R.S. (2010) Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res.*, **38**, 4274–4284.
44. Armitage,A.E., Deforche,K., Chang,C.H., Wee,E., Kramer,B., Welch,J.J., Gerstoft,J., Fugger,L., McMichael,A., Rambaut,A. *et al.* (2012) APOBEC3G-induced hypermutation of human immunodeficiency virus type-1 is typically a discrete 'all or nothing' phenomenon. *PLoS Genet.*, **8**, e1002550.
45. Armitage,A.E., Deforche,K., Welch,J.J., Van Laethem,K., Camacho,R., Rambaut,A. and Iversen,A.K. (2014) Possible footprints of APOBEC3F and/or other APOBEC3 deaminases, but not APOBEC3G, on HIV-1 from patients with acute/early and chronic infections. *J. Virol.*, **88**, 12882–12894.
46. Gillick,K., Pollpeter,D., Phalora,P., Kim,E.Y., Wolinsky,S.M. and Malim,M.H. (2013) Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4(+) T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination. *J. Virol.*, **87**, 1508–1517.
47. Kieffer,T.L., Kwon,P., Nettles,R.E., Han,Y., Ray,S.C. and Siliciano,R.F. (2005) G–>A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *J. Virol.*, **79**, 1975–1980.
48. Kim,E.Y., Lorenzo-Redondo,R., Little,S.J., Chung,Y.S., Phalora,P.K., Maljkovic Berry,I., Archer,J., Penugonda,S., Fischer,W., Richman,D.D. *et al.* (2014) Human APOBEC3 induced mutation of human immunodeficiency virus type-1 contributes to adaptation and evolution in natural infection. *PLoS Pathog.*, **10**, e1004281.
49. Sato,K., Takeuchi,J.S., Misawa,N., Izumi,T., Kobayashi,T., Kimura,Y., Iwami,S., Takaori-Kondo,A., Hu,W.S., Aihara,K. *et al.* (2014) APOBEC3D and APOBEC3F potently promote HIV-1 diversification and evolution in humanized mouse model. *PLoS Pathog.*, **10**, e1004453.
50. Land,A.M., Ball,T.B., Luo,M., Pilon,R., Sandstrom,P., Embree,J.E., Wachihi,C., Kimani,J. and Plummer,F.A. (2008) Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. *J. Virol.*, **82**, 8172–8182.
51. Wichroski,M.J., Robb,G.B. and Rana,T.M. (2006) Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. *PLoS Pathog.*, **2**, e41.
52. Ara,A., Love,R.P. and Chelico,L. (2014) Different mutagenic potential of HIV-1 restriction factors APOBEC3G and APOBEC3F is determined by distinct single-stranded DNA scanning mechanisms. *PLoS Pathog.*, **10**, e1004024.
53. Feng,Y., Love,R.P., Ara,A., Baig,T.T., Adolph,M.B. and Chelico,L. (2015) Natural polymorphisms and oligomerization of human APOBEC3H contribute to single-stranded DNA scanning ability. *J. Biol. Chem.*, **290**, 27188–27203.
54. Unutmaz,D., KewalRamani,V.N., Marmon,S. and Littman,D.R. (1999) Cytokine signals are sufficient for HIV-1 infection of resting human T lymphocytes. *J. Exp. Med.*, **189**, 1735–1746.
55. Yee,J.K., Friedmann,T. and Burns,J.C. (1994) Generation of high-titer pseudotyped retroviral vectors with very broad host range. *Methods Cell Biol.*, **43 Pt A**, 99–112.
56. Boussif,O., Lezoualc'h,F., Zanta,M.A., Mergny,M.D., Scherman,D., Demeneix,B. and Behr,J.P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 7297–7301.
57. Chen,J., Nikolaitchik,O., Singh,J., Wright,A., Bencsics,C.E., Coffin,J.M., Ni,N., Lockett,S., Pathak,V.K. and Hu,W.S. (2009) High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 13535–13540.
58. Burdick,R.C., Hu,W.S. and Pathak,V.K. (2013) Nuclear import of APOBEC3F-labeled HIV-1 preintegration complexes. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E4780–E4789.
59. Zenklusen,D., Larson,D.R. and Singer,R.H. (2008) Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.*, **15**, 1263–1271.
60. OhAinle,M., Kerns,J.A., Li,M.M., Malik,H.S. and Emerman,M. (2008) Antiretroelement activity of APOBEC3H was lost twice in recent human evolution. *Cell Host Microbe*, **4**, 249–259.
61. Kearney,M.F., Spindler,J., Shao,W., Yu,S., Anderson,E.M., O'Shea,A., Rehm,C., Poethke,C., Kovacs,N., Mellors,J.W. *et al.*

(2014) Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy. *PLoS Pathog.*, **10**, e1004010.
62. Palmer,S., Kearney,M., Maldarelli,F., Halvas,E.K., Bixby,C.J., Bazmi,H., Rock,D., Falloon,J., Davey,R.T. Jr, Dewar,R.L. *et al.* (2005) Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J. Clin. Microbiol.*, **43**, 406–413.
63. Agresti,A. (1990) *Categorical Data Analysis*. 2nd edn. Wiley, NY.
64. Team,C. (2012) *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
65. Burdick,R., Smith,J.L., Chaipan,C., Friew,Y., Chen,J., Venkatachari,N.J., Delviks-Frankenberry,K.A., Hu,W.S. and Pathak,V.K. (2010) P body-associated protein Mov10 inhibits HIV-1 replication at multiple stages. *J. Virol.*, **84**, 10241–10253.
66. Wang,X., Li,X., Ma,J., Zhang,L., Ma,L., Mi,Z., Zhou,J., Guo,F., Kleiman,L. and Cen,S. (2014) Human APOBEC3F incorporation into human immunodeficiency virus type 1 particles. *Virus Res.*, **191**, 30–38.
67. Bishop,K.N., Holmes,R.K., Sheehy,A.M., Davidson,N.O., Cho,S.J. and Malim,M.H. (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol.*, **14**, 1392–1396.
68. Ooms,M., Brayton,B., Letko,M., Maio,S.M., Pilcher,C.D., Hecht,F.M., Barbour,J.D. and Simon,V. (2013) HIV-1 Vif adaptation to human APOBEC3H haplotypes. *Cell Host Microbe*, **14**, 411–421.
69. Liddament,M.T., Brown,W.L., Schumacher,A.J. and Harris,R.S. (2004) APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. *Curr. Biol.*, **14**, 1385–1391.
70. Mbisa,J.L., Barr,R., Thomas,J.A., Vandegraaff,N., Dorweiler,I.J., Svarovskaia,E.S., Brown,W.L., Mansky,L.M., Gorelick,R.J., Harris,R.S. *et al.* (2007) Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. *J. Virol.*, **81**, 7099–7110.
71. Li,X.Y., Guo,F., Zhang,L., Kleiman,L. and Cen,S. (2007) APOBEC3G inhibits DNA strand transfer during HIV-1 reverse transcription. *J. Biol. Chem*, **282**, 32065–32074.
72. Iwatani,Y., Chan,D.S., Wang,F., Maynard,K.S., Sugiura,W., Gronenborn,A.M., Rouzina,I., Williams,M.C., Musier-Forsyth,K. and Levin,J.G. (2007) Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res.*, **35**, 7096–7108.
73. Holmes,R.K., Koning,F.A., Bishop,K.N. and Malim,M.H. (2007) APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J. Biol. Chem.*, **282**, 2587–2595.
74. Yu,Q., Konig,R., Pillai,S., Chiles,K., Kearney,M., Palmer,S., Richman,D., Coffin,J.M. and Landau,N.R. (2004) Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.*, **11**, 435–442.
75. Senavirathne,G., Jaszczur,M., Auerbach,P.A., Upton,T.G., Chelico,L., Goodman,M.F. and Rueda,D. (2012) Single-stranded DNA scanning and deamination by APOBEC3G cytidine deaminase at single molecule resolution. *J. Biol. Chem.*, **287**, 15826–15835.
76. Losey,H.C., Ruthenburg,A.J. and Verdine,G.L. (2006) Crystal structure of Staphylococcus aureus tRNA adenosine deaminase TadA in complex with RNA. *Nat. Struct. Mol. Biol.*, **13**, 153–159.
77. Janini,M., Rogers,M., Birx,D.R. and McCutchan,F.E. (2001) Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. *J. Virol.*, **75**, 7973–7986.
78. Li,Y., Kappes,J.C., Conway,J.A., Price,R.W., Shaw,G.M. and Hahn,B.H. (1991) Molecular characterization of human immunodeficiency virus type 1 cloned directly from uncultured human brain tissue: identification of replication-competent and -defective viral genomes. *J. Virol.*, **65**, 3973–3985.
79. Vartanian,J.P., Henry,M. and Wain-Hobson,S. (2002) Sustained G–>A hypermutation during reverse transcription of an entire human immunodeficiency virus type 1 strain Vau group O genome. *J. Gen. Virol.*, **83**, 801–805.
80. Russell,R.A., Moore,M.D., Hu,W.S. and Pathak,V.K. (2009) APOBEC3G induces a hypermutation gradient: purifying selection at multiple steps during HIV-1 replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA. *Retrovirology*, **6**, 16.
81. Refsland,E.W., Hultquist,J.F., Luengas,E.M., Ikeda,T., Shaban,N.M., Law,E.K., Brown,W.L., Reilly,C., Emerman,M. and Harris,R.S. (2014) Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity. *PLoS Genet.*, **10**, e1004761.
82. Bishop,K.N., Verma,M., Kim,E.Y., Wolinsky,S.M. and Malim,M.H. (2008) APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog.*, **4**, e1000231.
83. Mitra,M., Singer,D., Mano,Y., Hritz,J., Nam,G., Gorelick,R.J., Byeon,I.J., Gronenborn,A.M., Iwatani,Y. and Levin,J.G. (2015) Sequence and structural determinants of human APOBEC3H deaminase and anti-HIV-1 activities. *Retrovirology*, **12**, 3.
84. Delviks-Frankenberry,K.A., Nikolaitchik,O.A., Burdick,R.C., Gorelick,R.J., Keele,B.F., Hu,W.S. and Pathak,V.K. (2016) Minimal contribution of APOBEC3-Induced G-to-A hypermutation to HIV-1 recombination and genetic variation. *PLoS Pathog.*, **12**, e1005646.
85. Hersperger,A.R., Migueles,S.A., Betts,M.R. and Connors,M. (2011) Qualitative features of the HIV-specific CD8+ T-cell response associated with immunologic control. *Curr. Opin. HIV AIDS*, **6**, 169–173.
86. Goulder,P.J. and Watkins,D.I. (2008) Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat. Rev. Immunol.*, **8**, 619–630.
87. Casartelli,N., Guivel-Benhassine,F., Bouziat,R., Brandler,S., Schwartz,O. and Moris,A. (2010) The antiviral factor APOBEC3G improves CTL recognition of cultured HIV-infected T cells. *J. Exp. Med.*, **207**, 39–49.
88. Norman,J.M., Mashiba,M., McNamara,L.A., Onafuwa-Nuga,A., Chiari-Fort,E., Shen,W. and Collins,K.L. (2011) The antiviral factor APOBEC3G enhances the recognition of HIV-infected primary T cells by natural killer cells. *Nat. Immunol.*, **12**, 975–983.
89. Monajemi,M., Woodworth,C.F., Benkaroun,J., Grant,M. and Larijani,M. (2012) Emerging complexities of APOBEC3G action on immunity and viral fitness during HIV infection and treatment. *Retrovirology*, **9**, 35.
90. Josefsson,L., King,M.S., Makitalo,B., Brannstrom,J., Shao,W., Maldarelli,F., Kearney,M.F., Hu,W.S., Chen,J., Gaines,H. *et al.* (2011) Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 11199–11204.
91. Cha,L., Berry,C.M., Nolan,D., Castley,A., Fernandez,S. and French,M.A. (2014) Interferon-alpha, immune activation and immune dysfunction in treated HIV infection. *Clin. Transl. Immunology*, **3**, e10.
92. Hardy,G.A., Sieg,S., Rodriguez,B., Anthony,D., Asaad,R., Jiang,W., Mudd,J., Schacker,T., Funderburg,N.T., Pilch-Cooper,H.A. *et al.* (2013) Interferon-alpha is the primary plasma type-I IFN in HIV-1 infection and correlates with immune activation and disease markers. *PloS One*, **8**, e56527.
93. Raposo,R.A., Abdel-Mohsen,M., Bilska,M., Montefiori,D.C., Nixon,D.F. and Pillai,S.K. (2013) Effects of cellular activation on anti-HIV-1 restriction factor expression profile in primary cells. *J. Virol.*, **87**, 11924–11929.
94. Mohanram,V., Skold,A.E., Bachle,S.M., Pathak,S.K. and Spetz,A.L. (2013) IFN-alpha induces APOBEC3G, F, and A in immature dendritic cells and limits HIV-1 spread to CD4+ T cells. *J. Immunol.*, **190**, 3346–3353.
95. Stavrou,S. and Ross,S.R. (2015) APOBEC3 proteins in viral immunity. *J. Immunol.*, **195**, 4565–4570.