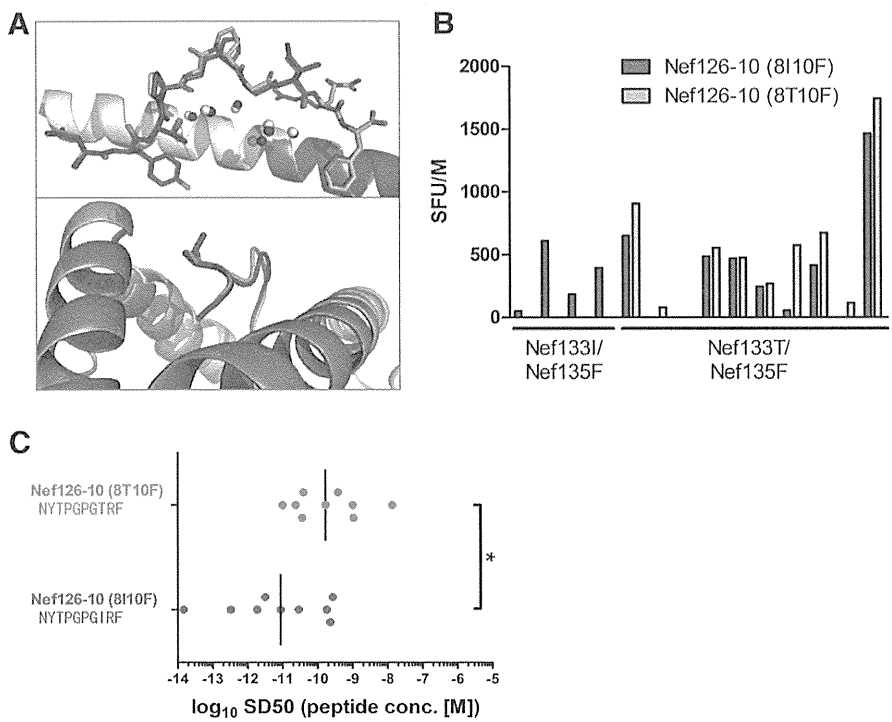


Figure 6 Impact of I133T mutation on the structures of pHLA and responding T cell subsets. (A) Top; Side view of structures of the Nef126-10(8I10F) (colored purple) and Nef126-10(8T10F) (colored green) is shown. Superimposition is based on the HLA-A*24:02 peptide binding domain. Water molecules are represented in blue and yellow spheres for the Nef126-10(8I10F) and Nef126-10(8T10F) epitopes, respectively. The number of peptide residues is labeled in black through P1 to P10. The HLA-A*24:02 is represented as a dark grey (Nef126-10(8T10F) and light grey cartoon. **(B)** The presence of the responses to each peptide was examined by *ex vivo* ELISpot assay. A pair of sticks, Nef126-10(8I10F) (purple) and Nef126-10(8T10F) (green), represents the average SFU of duplicate wells to each peptide in one patient. Four individuals on the left side harbored 133I/135F mutant and 10 individuals on the right side harbored 133T/135F mutant. **(C)** Functional avidity curve was drawn for 9 subjects who harbored Nef133T/135F sequence. From the curve, SD50 was calculated and projected in this chart. Nef126-10(8I10F)-specific CTL responses (purple) showed significantly higher functional avidity compared to Nef126-10(8T10F) (green). Each point represents functional avidity for specific peptide in one individual. The significance between two groups was calculated by Mann–Whitney U-test; *, $p < 0.05$.

persons harboring 135F is consistent with escape at this position affording clear viral advantage in at least some cases.

We therefore propose the following model of HLA-A*24:02-mediated CTL targeting of these overlapping Nef epitopes *in vivo*. If an HLA-A*24:02-expressing patient were to be infected with the wild type virus, CTL responses would first arise against Nef134-10, eliciting the Y135F escape mutant. This in turn would reveal the novel Nef126-10 epitope, against which CTL responses would then arise. Similarly, if an HLA-A*24:02 expressing patient were to be infected with a virus harboring Y135F (a likely occurrence given its extremely high prevalence in Japan), we infer that a CTL response against Nef126-10 epitope would be launched immediately following infection.

As such HIV's evasion of Nef134-10-specific CTL (either via transmission or *in vivo* selection of Y135F) yields a major, albeit temporary advantage to HIV, that is subsequently diminished by the creation of the Nef126-10 epitope that is then targeted in most A*24:02-expressing persons. By extension, the "epitope switching" from Nef134-10 to Nef126-10 may contribute in part to the previous observation of relatively stable pVL over time among HLA-A*24:02 positive patients [26] despite high transmission frequency of 135F [4]. These observations are not inconsistent with the idea that population-level HIV adaptation to HLA could lead to weakening of host antiviral control if "HLA-adapted" forms dominated the population (as is the case with Nef-135F in Japan, which represents the consensus at this position) [27]. If the consensus at Nef codon 135 was the susceptible Y rather than the escaped F in Japan, we hypothesize that control of HIV by A*24 would be even greater in this region, as Nef134–10 and Nef126-10 would be sequentially targeted (rather than just Nef126-10 in the case where 135F is acquired at transmission).

Of note, among the 46 patients in the IMSUT cohort, there were no significant differences in plasma HIV load between patients with or without Nef126-10-specific

CD8+ T cell responses (Figure 1, and data not shown). However, this observation should be interpreted in the context that all patients responding to Nef126-10 responses harbored Nef-135F (that confers escape from responses against the downstream Nef134-10 epitope). Moreover, the observation that no reversion of 135F to 135Y was ever observed in Nef126-10 responders (Figure 1C) suggests that Nef126-10-specific CTL responses are not as effective as Nef134-10-specific ones in controlling HIV *in vivo*. Elucidating the mechanisms and *in vivo* relevance of our observations would be crucial for vaccine development.

Among the HLA-associated polymorphisms, amino acid substitutions between K-R, E-D, V-I, I-L, and Y-F are relatively frequent [8,24,25]. The similarities of these amino acid residue pairs with respect to size, charge, hydrophobicity, and other biochemical properties suggest that they are often critical to the structure and function of the viral proteins involved. Although in the present study we characterized an example of escape-induced "epitope switching" in HIV-1, we do not know how often this phenomenon occurs. If an HLA-driven escape mutation has the potential to serve as an anchor of a cryptic epitope, and if an N-terminal or C-terminal anchor residue is present at a proper distance, incorporation of both the wild-type and the mutant peptide in a vaccine may elicit immune response to that cryptic epitope. The vaccine-induced CTL repertoire may be useful following selection of the escape mutant or may prevent its selection altogether. As such, the phenomenon of escape-induced "epitope switch" may be relevant to vaccine design.

Conclusions

Our data represent the first example of the *de novo* creation of a novel overlapping CTL epitope as a direct result of HLA-driven immune escape in a neighboring epitope. The robust targeting of Nef126-10 following transmission (or *in vivo* selection) of HIV-1 containing Y135F may explain in part the previously reported stable plasma viral loads over time in the Japanese population, despite the high prevalence of both HLA-A*24:02 and Nef-Y135F in circulating HIV-1 sequences.

Methods

Patients and samples

The samples and host/viral genotype data analyzed in this study were obtained from three independent sources: (i) the IMSUT cohort at the Institute of Medical Science, the University of Tokyo, Japan, (ii) the baseline (pre-therapy) cross-section of the HOMER cohort in British Columbia, Canada ([8] and unpublished), and (iii) a longitudinal multicenter cohort of acute/early infected individuals [28]. The IMSUT cohort, which consists primarily of Asian patients with chronic infection, was used to determine the viral sequences and immune responses.

IMSUT cohort

Forty-six HLA-A*24:02-positive, antiretroviral-naïve, chronically HIV-infected subjects were selected from among patients participating in an ongoing HIV-1-immunopathogenesis study at an HIV outpatient clinic affiliated with the Institute of Medical Science, the University of Tokyo (IMSUT). Study procedures included routine collection of blood samples for virologic and immunologic testing. Peripheral blood mononuclear cells (PBMCs) and plasma samples were separated and preserved in liquid nitrogen or at −80°C, respectively, until use. The study was approved by the internal review board of the Institute of the Medical Science of the University of Tokyo (No. 11–2), and all subjects were adults and provided written informed consent.

HOMER cohort

A total of 1038 patients from the HAART Observational Medical Evaluation and Research (HOMER) cohort, an open cohort of initially antiretroviral-naïve chronically HIV-infected individuals in British Columbia, Canada of predominantly Caucasian ethnicity, were analyzed in the present study. Plasma HIV-1 RNA sequencing and HLA class I sequence-based typing were performed as previously described [8]. We applied phylogenetically-corrected methods [25] to determine the strength of association between amino acid variants at Nef codons 133 and 135 in this dataset.

Longitudinal acute/early infection cohort

Kaplan-Meier analysis was used to investigate the time course of selection of specific immune escape mutations at Nef codons 133 and 135 among 16 HLA-A*24 expressing individuals from a longitudinal, multicenter, acute/early HIV-1 infection cohort [28]. "Time to escape" was defined as the number of days elapsed between estimated infection date and first detection of the escape variant (as a full or partial amino acid change).

Plasma viral RNA sequences

Viral RNA was extracted from 140 μl of plasma using the QIAamp viral RNA Mini kit (QIAGEN). Using 4 μl of RNA as starting material, reverse transcription and first PCR were carried out according to the manufacturer's protocol with SuperScript III One-Step RT-PCR System with Platinum Taq High Fidelity (Invitrogen). Two μl of the first PCR product was subjected to nested PCR, performed using Ex-Taq HS (Takara) with 35 cycles of 30s at 94°C, 30s at 55°C, 60s at 72°C and a final extension for 7 min at 72°C. The primer sets were as follows (Nucleotide positions are those of the published HIV-1 SF2 strain (GenBank accession number: K02007). For the first PCR, primers Nef-1F (5'-GTAGCTGAGGGGACAGATAGGGTTAT-3') (nt 8,6

88 to 8,731) and Nef-1R (5'-GCACTCAAGGCAAGCT TTATTGAGGC-3') (nt 9,632 to 9,607) were used; and for the nested PCR, primers Nef-2F (5'-CGTCTAGAACAT ACCTAGAAGAATAAGACAGG-3') (nt 8,746 to 8,777) and Nef-2R (5'-CGGAATCCGTCCCCGCGGAAAGTCCC TTGTA-3') (nt 9,477 to 9,444) were used. The PCR products were purified with a PCR purification kit (QIAGEN) before sequencing. DNA sequencing was performed using an ABI Prism dye terminator cycle sequencing-ready reaction kit (Applied Biosystems) on a Perkin-Elmer ABI-377 sequencer.

Expression vectors

To construct the HLA-A*24:02 expression vector, pcDNA3.1-A24-DsRedm, the HLA-A*24:02 sequence was amplified by PCR using cloned HLA-A*24:02 cDNA as a template [29,30] and digested with BamHI and NcoI. Primers (5'-TAATACGACTCACTATAGGG-3') and (5'-CCATGGATCCGCCCCCTCCCACTTTACAAGCTGT GAGAGACAC-3') were used for the amplification. The pDsRed-Monomer (Clontech) sequence was digested by NcoI and NotI, and was ligated to the 3' end of HLA-A*24:02 fragment to obtain the HLA-A24-DsRedm fragment. Then HLA-A24-DsRedm fragment was inserted into the multiple cloning site of pcDNA3.1/Hygro(+) vector (Invitrogen).

Mini-Nef and mini-Gag gene expression vectors containing two reporter genes, renilla luceferase (Rluc) and EGFP, together with hygromycin selection were constructed in a pTracer-CMV2 vector (Invitrogen) as follows. Wild-type and mutant mini-Nef genes (from amino acid position 123 to 153) were amplified by PCR using the plasmids containing HIV-1 SF-2 Nef gene with wild-type or mutant sequence as a template [4,31,32]. Primer sequences were (5'-GGTACCGCCGCCATGGATTGG-CAGAATTACACA-3') and (5'-GGATCCGCCCCCTC CTACCTTCTCTGGCTC-3'). A mini-Gag gene, extending from amino acid position 18 to 46 in p17, which includes the HLA-A*24:02-restricted CTL epitope Gag28-9 (KYRLKHIVW), was amplified using a plasmid containing the 5' half of the HIV-1 SF2 strain [31,32]. The primers used were 5'-GGTACCGCCGCCATGAAAAT TCGGTTAAGG-3' and 5'-GGATCCGCCCCCTCCGA CTGCGAATCGTTC-3'. The Rluc and hygromycin genes were amplified from pGL4.77 hRlucP/Hygro (Promega) using primers 5'-GGATCCATGGCTTCCAAGGTGTA C-3' and 5'-TCTAGAGTCGCGGCCTTAGACGTT-3'.

KpnI/BamHI fragments of mini-Nef or mini-Gag, BamHI/XbaI fragment of Rluc and hygromycin genes amplified from pGL4.77 hRlucP/Hygro were ligated to the KpnI/XbaI fragment of pTracer-CMV2 vector (Invitrogen) to create a mini-Nef or mini-Gag expression vector, pmNef(wt)-hRluc-EGFP and pmGag(wt)-hRluc-EGFP, respectively. In the final step the GFPz sequence was replaced by EGFP sequence (Clontech).

Peptides

Synthetic peptides were purchased from Sigma-Genosys. The peptides used in the screening of immune response by ELISpot had a purity of 70% or more. All other peptides were more than 95% pure as determined by high-performance liquid chromatography and mass spectroscopy.

Cells and media

T2-A24, a kind gift from K. Kuzushima, was cultured in RPMI 1640 (Sigma) supplemented with 100 U of penicillin/ml, 100 U of streptomycin/ml, 10% heat-inactivated fetal calf serum (FCS) (Sigma), and 0.8 mg of G418 (Invitrogen)/ml [33]. We established Nef126-10 and Nef134-10-specific CTL clones, I30-1 and H27-9, as previously described [18]. CTL clones were cultured with RPMI 1640 supplemented with 50 U of interleukin-2/ml, 100U of penicillin/ml, 100U of streptomycin/ml, and 10% heat-inactivated FCS (R10/50), but the clones were cultured in the absence of interleukin-2 (R10) for two days before antigen presentation assays. pcDNA3.1-A24-DsRedm was introduced into 293FT cell line (Invitrogen) and the cells were treated by hygromycin for 2 weeks. After cloning by limiting dilution we obtained 293FT-A24DRm-CY0, and confirmed HLA-A*24:02 expression with FACS analysis by using anti-HLA-A9 serotype antibody (One Lambda, data not shown).

IFN-γ ELISpot assay

The gamma interferon enzyme-linked immunospot (IFN-γ ELISpot) assay was performed using patients' PBMCs as previously described [4] with some modifications. In brief, 96-well plates (Millipore) were coated with anti-gamma-interferon (IFN-γ) MAb 1-D1k (Mabtech) overnight at 4°C. Peptides were added directly to the wells at a final concentration of 10^{-5} M. $5 \sim 10 \times 10^4$ cells were added to each well with a final volume of 100 μl of R10. For negative controls, PBMCs were incubated with R10 alone without peptides. After incubation at 37°C under 5% CO_2 overnight (16 to 18 h), the plates were washed six times with phosphate-buffered saline containing 0.01% tween-20 (PBST). Biotinylated anti-IFN-γ MAb 7-B6-1 (Mabtech) was added, and was incubated for 2 hours at 37°C under 5% CO_2. After washing with PBST, streptavidin-alkaline phosphatase conjugate (Mabtech) was added and the plates were kept at room temperature for 45 min. After washing with PBST, IFN-γ-producing cells were detected as dark spots after 10- to 20-min color reaction with 5-bromo-4-chloro-3-indolylphosphate and nitroblue tetrazolium by using AP Conjugate Substrate Kit (Bio-Rad). Spots were counted by KS ELISPOT compact (Carl Zeiss) and expressed as spot-forming units (SFU) per 10^6 PBMCs

after subtracting the SFU of the negative control. Values with >50 SFU, >3 × mean SFU of negative control and > mean SFU of negative control + 3 SD per 10^6 input cells were considered as a positive response.

Since more cells were required for the immune response screening (Figure 1B) and functional avidity assays (Figures 4B and 6C), PBMCs were stimulated with anti-human CD3 antibody and the T cells were expanded for 2 to 3 weeks in R10/50 (BD Pharmingen). Culture media was changed to R10 two days prior to the assay date. For *ex vivo* IFN-γ ELISpot assay (Figures 5B and 6B), PBMCs were cultured for 6 hours in R10 media.

Peptide-HLA binding assay (Figure 2A)

Peptide binding to HLA-A*24:02 was assessed by using a T2-A24 stabilization assay as previously described [4,33]. Briefly, after incubation for 16 hours at 26°C under 5% CO_2, 2×10^5 T2-A24 cells were incubated with 10^{-4} to 10^{-9} M peptides for 1 h at 4°C. After keeping at 37°C under 5% CO_2 for 3 hours, the cells were stained with biotinylated anti-human HLA-A9 monoclonal antibody (One Lambda), and streptavidin-APC conjugates (BD Pharmingen). The mean fluorescence intensity (MFI) was measured by FACSCalibur (Becton Dickinson). In each experiment, MFI of samples was normalized by the MFI of 10^{-4} M control peptide, Nef134-8(RYPLTFGW). Three independent experiments were performed.

CTL clones and Epitope recognition (Figure 2B)

Nef134-10- and Nef126-10-specific clones, H27-9 and I30-1, were established by Nef134-10(wt) and Nef126-10 (wt) peptide stimulation respectively, and limiting dilution of PBMCs from HIV-1-infected patients harboring 133 T/135 F. For *in-vitro* peptide stimulation, 5×10^5 PBMCs were pulsed with 10 μM of each peptide for 1 h. The cells were washed twice with R10, then cultured with 1×10^6 autologous PBMCs and 4×10^6 irradiated (3300 rad) allogeneic PBMCs in R10. After 4 days, IL-2 was added to 50 U/ml and the cells were cultured for 2–3 weeks in R10/50. Peptide-specific $CD8^+$ T cells were enriched by MACS separation (Miltenyi) using tetramers. The sorted cells were cloned by limiting dilution to 3 or 10 cells/well in 96-well round-bottom tissue culture plates, and cultured with 10^5 irradiated allogeneic PBMCs in R10/50 containing 5 μg/ml PHA-L.

CTL-recognition of the epitopes was assessed by serially diluted peptides. On day 0, 293FT-A24DRm-CY0 cells were seeded onto 96-well Flat-bottom transparent plate (BD Falcon) so that the cultures become confluent on day 2. On day 2, each peptide was pulsed with concentrations from 3^2 to 3^{-6} μM to the wells and incubated at 37°C under 5% CO_2 for 1 h. Then CTLs (5,000 ~ 10,000 cells) were added and co-cultured at 37°C under 5% CO_2 for 18 to 24 h. After the incubation, supernatants were harvested, and IFN-γ concentrations were quantified by Human IFN-γ ELISA Set (BD Bioscience). In each experiment, the IFN-γ value of samples was normalized to that of the highest wild type peptide concentration (9 μM). For example, each IFN-γ value of I30-1 CTL clone was divided by the value of the well pulsed with 9 μM Nef126-10(wt) peptide. Each assay was performed in duplicate and three independent experiments were conducted.

Antigen presentation assay (Figure 3)

Antigen presentation was assessed by measuring epitope-specific CTL responses to endogenously expressed antigen. First, intracellular expression of each antigen was extrapolated from the activity of reporter protein, Rluc. 293FT-A24DRm-CY0 cells were seeded in 96-well plate (Nunc) on day 0. On day 1, antigen expression plasmids, pmNef(wt)-hRluc-EGFP, pmNef(135F)-hRluc-EGFP, pmNef (133T135F)-hRluc-EGFP, or control vectors (pmGag(wt)-hRluc-EGFP) were transfected into the cells in each well using FuGENE HD (Promega). The cultures were incubated at 37°C under 5% CO_2 for 18 to 24 h. On day 2, transfection efficiency was inspected roughly under a fluorescence microscope (KEYENCE BZ-9000), then Rluc activity in each transfected well was measured by using Dual-Glo Luciferase Assay System (Promega) and luminometer (Promega GloMax 96 Microplate Luminometer). The assay was performed in triplicate.

Second, CTL responses against endogenously expressed and processed epitopes were evaluated. 293FT-A24DRm-CY0 cells were seeded onto a 96-well Flat-bottom transparent plate (BD Falcon) on day 0. On day 1, expression plasmid was transfected to each well using FuGENE HD (Promega). CTLs (5,000 ~ 10,000 cells) were added to the transfected wells on day 2. After incubation for 18 to 24 h, the supernatant of each well was harvested, and IFN-γ secretion was quantified by Human IFN-γ ELISA Set (BD Bioscience). IFN-γ ELISA was performed in duplicate.

Experiments were performed in duplicate on three independent occasions. To normalize the values in each experiment, mean IFN-γ values of each sample were normalized by the mean value of reference wells in duplicate. In the reference wells, 9 μM of the wild type peptide was pulsed to the antigen presenting cells and then co-cultured with CTLs.

Protein expression in E. coli, refolding and purification

HLA-A*24:02 and β2m were expressed in *E. coli* and refolded from inclusion bodies as previously described with some modifications [34]. HLA-A*24:02 (18 mg), β2m (6 mg) and peptide (4 mg) were mixed in 400 ml of refold buffer containing 100 mM Tris, pH 8.0, 400 mM L-arginine-HCl, 2 mM EDTA, 5 mM GSH, 0.5 mM GSSG, 0.2 mM PMSF. The refolded protein was purified by Superdex 75 column, followed by Mono Q column, and

subsequently concentrated to 10 mg/ml in 20 mM Tris, pH 8.0, 50 mM NaCl for crystallization.

Crystallization, data collection and structure determination

The crystallization was done by the sitting drop vapor diffusion method at 20°C. Crystals of the A24/N126-10 (8T10F) complex were obtained in 20% (w/v) PEG 3350, 200 mM sodium phosphate dibasic, and those of A24/N126-10(8I10F) were obtained in 20% (w/v) PEG 3350, 200 mM sodium nitrate. For cryoprotection, crystals were soaked briefly in reservoir solutions containing 20% ethylene glycol, and then frozen in liquid nitrogen before data collection. Data were collected at the beamline BL41XU in SPring 8 (Hyogo, Japan), and processed with HKL2000 [35] and the CCP4 program suite [36].

The structure were determined by molecular replacement using Molrep [37]. The search model was the coordinate file of PDB (Protein Data Bank) code 3I6L with omitted peptide for A24/N126-10(8T10F). Model building and refinement were carried out using Coot [38] and REFMAC5.6 implemented in CCP4, respectively. The structure of A24/N126-10(8I10F) was determined with the refined A24/N126-10(8T10F) as a search model and refined as described above. The stereochemistry of the refined models was assessed with RAMPAGE [39]. All molecular graphic representations were created with the program PyMOL (DeLano Scientific; http://www.pymol.org). Data collection and refinement statistics are shown in Additional file 2: Table S1.

Functional avidity assay

Basically, the assay was performed using the same procedure with IFN-γ ELISpot assay as described above. PBMCs were cultured for 2 to 3 weeks in R10/50 after anti-human CD3 antibody (BD Pharmingen) stimulation, and culture media were changed from R10/50 to R10 two days before use. PBMCs were incubated with peptides at concentrations from 10^{-5} to 10^{-12} M, and SFU was calculated. The functional avidity to peptide dilutions was determined as a 50% of sigmoidal dose (SD_{50}) SFU.

Statistical analysis

All data visualization and statistical analyses were performed using GraphPad Prism (GraphPad Software, La Jolla, CA). Student's t-test and Mann–Whitney U-test were used to compare the antigen presentation and functional avidity between two groups, respectively. Spearman rank correlation was used to calculate the correlation between peptide-specific response and pVL. Dose at 50% response in sigmoidal dose–response curves (SD_{50}) was calculated by drawing sigmoidal dose–response curves. Time to mutational escape, defined as the time elapsed between estimated date of HIV-1 infection and the first appearance of a full or partial amino acid change consistent with the specific escape mutation of interest, was calculated using Kaplan-Meier (survival) methods.

Additional files

Additional file 1: Figure S1. Overview of structures of the HLA-A*2402 in complex with the Nef126-10 peptides. Structures of **(A)** the A24/N126-10 (8I10F) and **(B)** the A24/N126-10(8T10F).The electron density of **(C)** the Nef126-10(8I10F) and **(D)** the Nef126-10(8T10F) are shown with F_o-F_c omit maps contoured at 2.0 σ (cyan mesh). (**C** and **D**) The peptide structures are shown in a side view (top panels) and top view (bottom panels). The Nef126-10 (8I10F) and the Nef126-10 (8T10F) are shown as a purple and a green stick model, respectively. HLA-A24 and β2m are represented as gray and black cartoon model, respectively.

Additional file 2: Table S1. Data collection and refinement statistics.

Competing interests

AI. has received grant support from Toyama Chemical Co. Ltd., astellas, ViiV Healthcare KK., MSD KK., Baxter through the University of Tokyo. AI. has received speaker's honoraria/payment for manuscript from Eiken Chemical Co. Ltd., astellas, Toyama Chemical Co. Ltd, Torii Pharmaceutical Co. Ltd., MSD KK., and Taisho Toyama Pharmaceutical Co. Ltd. The authors have no additional financial interests.

Authors' contributions

CH conceived of the study, carried out the molecular genetic studies and immunoassays, and drafted the manuscript. AKT established CTL clones and construction of expression vector. AS, YS, AY and SF carried out protein synthesis and solved crystal structures. DZ contributed to the sequence analysis. HN, EA, TKi, MK and TKo provided the clinical materials and data of the IMSUT cohort. GFG participated to the study design. EM and ZLB investigated the relationship between HLA-A*24 expression and sequence variants at Nef codons 135 and 133 in 1018 participants in British Columbia HOMER cohort and longitudinal multicenter acute/early HIV-1 infection cohort with HIV-1 Nef and HLA-A data available. ZLB helped to draft the manuscript. AI participated in the study design, coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the beam-line staffs at NW12A and BL5A of Photon Factory (Tsukuba, Japan) and BL41XU of SPring8 (Hyogo, Japan) for technical help during data collection. We thank Drs. Richard Harrigan, Heiko Jessen, Anthony Kelleher, Martin Markowitz and Bruce Walker for specimen and/or data access.
This work was supported in part by a contract research fund from the Ministry of Education, Culture, Sports, Science and Technology(MEXT) for Program of Japan Initiative for Global Research Network on Infectious Diseases (10005010) (AI); Global COE Program (Center of Education and Research for Advanced Genome-Based Medicine - For personalized medicine and the control of worldwide infectious diseases) of MEXT (F06) (AI); Research on international cooperation in medical science, Research on global health issues, Health and Labour Science Research Grants, the Ministry of Health, Labor, and Welfare of Japan (H25-KOKUI-SITEI-001)(AI). Grants for AIDS research from the Ministry of Health, Labor, and Welfare of Japan (H24-AIDS-IPPAN-008) (AKT), H25-AIDS-IPPAN-007 (AKT); JSPS KAKENHI (25293226) (AKT); Research grant from Banyu Life Science Foundation International (AKT); Master's Scholarship from the Canadian Association of HIV Research and Abbott Virology (EM); CIHR New Investigator Award and a Scholar Award from the Michael Smith Foundation for Health Research (ZLB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Accession numbers

The coordinates and structure factors of the HLA-A24/Nef126-10(8I10F) and HLA-A24/Nef126-10(8T10F) have been deposited [Protein Data Bank: 3WL9 and 3WLB, respectively].

Author details
[1]Division of Infectious Diseases, Advanced Clinical Research Center, the Institute of Medical Science, the University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [2]Department of Infectious Disease Control, the International Research Center for Infectious Diseases, the Institute of Medical Science, the University of Tokyo, Tokyo, Japan. [3]Department of Infectious Diseases and Applied Immunology, Hospital, the Institute of Medical Science, the University of Tokyo, Tokyo, Japan. [4]CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China. [5]Structural Biology Laboratory, Life Science Division, Synchrotron Radiation Research Organization and Institute of Molecular and Cellular Biosciences, the University of Tokyo, Tokyo, Tokyo, Japan. [6]Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan. [7]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada. [8]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada. [9]Asian Research Center for Infectious Diseases, the Institute of Medical Science, the University of Tokyo, Tokyo, Japan. [10]Cancer Immunology Branch, Division of Cancer Biology, National Cancer Center, Goyang-si, Gyeongi-do 410-769, Korea.

Received: 14 January 2014 Accepted: 28 April 2014
Published: 21 May 2014

References

1. McMichael AJ, Rowland-Jones SL: **Cellular immune responses to HIV.** *Nature* 2001, **410**(6831):980–987.
2. Walker BD, Burton DR: **Toward an AIDS vaccine.** *Science* 2008, **320**(5877):760–764.
3. Kawana A, Tomiyama H, Takiguchi M, Shioda T, Nakamura T, Iwamoto A: **Accumulation of specific amino acid substitutions in HLA-B35-restricted human immunodeficiency virus type 1 cytotoxic T lymphocyte epitopes.** *AIDS Res Hum Retrovir* 1999, **15**(12):1099–1107.
4. Furutsuki T, Hosoya N, Kawana-Tachikawa A, Tomizawa M, Odawara T, Goto M, Kitamura Y, Nakamura T, Kelleher AD, Cooper DA, Iwamoto A: **Frequent transmission of cytotoxic-T-lymphocyte escape mutants of human immunodeficiency virus type 1 in the highly HLA-A24-positive Japanese population.** *J Virol* 2004, **78**(16):8437–8445.
5. Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, Elvin JG, Rothbard JA, Bangham CR, Rizza CR, McMichael A: **Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition.** *Nature* 1991, **354**(6353):453–459.
6. Geels MJ, Cornelissen M, Schuitemaker H, Anderson K, Kwa D, Maas J, Dekker JT, Baan E, Zorgdrager F, van den Burg R, van Beelen M, Lukashov VV, Fu TM, Paxton WA, van der Hoek L, Dubey SA, Shiver JW, Goudsmit J: **Identification of sequential viral escape mutants associated with altered T-cell responses in a human immunodeficiency virus type 1-infected individual.** *J Virol* 2003, **77**(23):12430–12440.
7. Allen TM, Yu XG, Kalife ET, Reyor LL, Lichterfeld M, John M, Cheng M, Allgaier RL, Mui S, Frahm N, Alter G, Brown NV, Johnston MN, Rosenberg ES, Mallal SA, Brander C, Walker BD, Altfeld M: **De novo generation of escape variant-specific CD8+ T-cell responses following cytotoxic T-lymphocyte escape in chronic human immunodeficiency virus type 1 infection.** *J Virol* 2005, **79**(20):12952–12960.
8. Brumme ZL, Brumme CJ, Heckerman D, Korber BT, Daniels M, Carlson J, Kadie C, Bhattacharya T, Chui C, Szinger J, Mo T, Hogg RS, Montaner JS, Frahm N, Brander C, Walker BD, Harrigan PR: **Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1.** *PLoS Pathog* 2007, **3**(7):e94.
9. Haas G, Plikat U, Debre P, Lucchiari M, Katlama C, Dudoit Y, Bonduelle O, Bauer M, Ihlenfeldt HG, Jung G, Maier B, Meyerhans A, Autran B: **Dynamics of viral variants in HIV-1 Nef and specific cytotoxic T lymphocytes in vivo.** *J Immunol* 1996, **157**(9):4212–4221.
10. Wibmer CK, Bhiman JN, Gray ES, Tumba N, Abdool Karim SS, Williamson C, Morris L, Moore PL: **Viral Escape from HIV-1 Neutralizing Antibodies Drives Increased Plasma Neutralization Breadth through Sequential Recognition of Multiple Epitopes and Immunotypes.** *PLoS Pathog* 2013, **9**(10):e1003738.
11. Itoh Y, Mizuki N, Shimada T, Azuma F, Itakura M, Kashiwase K, Kikkawa E, Kulski JK, Satake M, Inoko H: **High-throughput DNA typing of HLA-A, –B, –C, and -DRB1 loci by a PCR-SSOP-Luminex method in the Japanese population.** *Immunogenetics* 2005, **57**(10):717–729.
12. Fujiwara M, Tanuma J, Koizumi H, Kawashima Y, Honda K, Mastuoka-Aizawa S, Dohki S, Oka S, Takiguchi M: **Different abilities of escape mutant-specific cytotoxic T cells to suppress replication of escape mutant and wild-type human immunodeficiency virus type 1 in new hosts.** *J Virol* 2008, **82**(1):138–147.
13. Ikeda-Moore Y, Tomiyama H, Miwa K, Oka S, Iwamoto A, Kaneko Y, Takiguchi M: **Identification and characterization of multiple HLA-A24-restricted HIV-1 CTL epitopes: strong epitopes are derived from V regions of HIV-1.** *J Immunol* 1997, **159**(12):6242–6252.
14. Goulder PJ, Edwards A, Phillips RE, McMichael AJ: **Identification of a novel HLA-A24-restricted cytotoxic T-lymphocyte epitope within HIV-1 Nef.** *AIDS (London, England)* 1997, **11**(15):1883–1884.
15. Choppin J, Cohen W, Bianco A, Briand JP, Connan F, Dalod M, Guillet JG: **Characteristics of HIV-1 Nef regions containing multiple CD8+ T cell epitopes: wealth of HLA-binding motifs and sensitivity to proteasome degradation.** *J Immunol* 2001, **166**(10):6164–6169.
16. Ibe M, Moore YI, Miwa K, Kaneko Y, Yokota S, Takiguchi M: **Role of strong anchor residues in the effective binding of 10-mer and 11-mer peptides to HLA-A*2402 molecules.** *Immunogenetics* 1996, **44**(4):233–241.
17. Sidney J, Southwood S, Sette A: **Classification of A1- and A24-supertype molecules by analysis of their MHC-peptide binding repertoires.** *Immunogenetics* 2005, **57**(6):393–408.
18. Miyazaki E, Kawana Tachikawa A, Tomizawa M, Nunoya J, Odawara T, Fujii T, Shi Y, Gao GF, Iwamoto A: **Highly restricted T-cell receptor repertoire in the CD8+ T-cell response against an HIV-1 epitope with a stereotypic amino acid substitution.** *AIDS (London, England)* 2009, **23**(6):651–660.
19. Almeida JR, Sauce D, Price DA, Papagno L, Shin SY, Moris A, Larsen M, Pancino G, Douek DC, Autran B, Saez-Cirion A, Appay V: **Antigen sensitivity is a major determinant of CD8+ T-cell polyfunctionality and HIV-suppressive activity.** *Blood* 2009, **113**(25):6351–6360.
20. Mothe B, Llano A, Ibarrondo J, Zamarreno J, Schiaulini M, Miranda C, Ruiz-Riol M, Berger CT, Herrero MJ, Palou E, Plana M, Rolland M, Khatri A, Heckerman D, Pereyra F, Walker BD, Weiner D, Paredes R, Clotet B, Felber BK, Pavlakis GN, Mullins JI, Brander C: **CTL responses of high functional avidity and broad variant cross-reactivity are associated with HIV control.** *PLoS One* 2012, **7**(1):e29717.
21. Carlson JM, Brumme CJ, Martin E, Listgarten J, Brockman MA, Le AQ, Chui CK, Cotton LA, Knapp DJ, Riddler SA, Haubrich R, Nelson G, Pfeifer N, Deziel CE, Heckerman D, Apps R, Carrington M, Mallal S, Harrigan PR, John M, Brumme ZL, International HIV Adaptation Collaborative: **Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1.** *J Virol* 2012, **86**(24):13202–13216.
22. Chikata T, Carlson JM, Tamura Y, Borghan MA, Naruto T, Hashimoto M, Murakoshi H, Le AQ, Mallal S, John M, Gatanaga H, Oka S, Brumme ZL, Tachiguchi M: **Host-specific adaptation of HIV-1 subtype B in the Japanese population.** *J Virol* 2014, **88**(9):4764–4775.
23. Shimizu A, Kawana-Tachikawa A, Yamagata A, Han C, Zhu D, Sato Y, Nakamura H, Koibuchi T, Carlson J, Martin E, Brumme CJ, Shi Y, Gao GF, Brumme ZL, Fukai S, Iwamoto A: **Structure of TCR and antigen complexes at an immunodominant CTL epitope in HIV-1 infection.** *Sci Rep* 2013, **3**:3097.
24. Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA, Swenson LC, Tao I, Szeto S, Rosato P, Sela J, Kadie CM, Frahm N, Brander C, Haas DW, Riddler SA, Haubrich R, Walker BD, Harrigan PR, Heckerman D, Mallal S: **HLA-associated immune escape pathways in HIV-1 subtype B Gag, Pol and Nef proteins** *PloS one* 2009, **4**(8):e6687.
25. Carlson JM, Listgarten J, Pfeifer N, Tan V, Kadie C, Walker BD, Ndung'u T, Shapiro R, Frater J, Brumme ZL, Goulder PJ, Heckerman D: **Widespread impact of HLA restriction on immune control and escape pathways of HIV-1.** *J Virol* 2012, **86**(9):5230–5243.
26. Koga M, Kawana-Tachikawa A, Heckerman D, Odawara T, Nakamura H, Koibuchi T, Fujii T, Miura T, Iwamoto A: **Changes in impact of HLA class I allele expression on HIV-1 plasma virus loads at a population level over time.** *Microbiol Immunol* 2010, **54**(4):196–205.
27. Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA: **Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level.** *Science* 2002, **296**(5572):1439–1443.
28. Brumme ZL, Brumme CJ, Carlson J, Streeck H, John M, Eichbaum Q, Block BL, Baker B, Kadie C, Markowitz M, Jessen H, Kelleher AD, Rosenberg E, Kaldor J, Yuki Y, Carrington M, Allen TM, Mallal S, Altfeld M, Heckerman D,

Walker BD: **Marked epitope- and allele-specific differences in rates of mutation in human immunodeficiency type 1 (HIV-1) Gag, Pol, and Nef cytotoxic T-lymphocyte epitopes in acute/early HIV-1 infection.** *J Virol* 2008, **82**(18):9216–9227.
29. Tokunaga K, Ishikawa Y, Ogawa A, Wang H, Mitsunaga S, Moriyama S, Lin L, Bannai M, Watanabe Y, Kashiwase K, Tanaka H, Akaza T, Tadokoro K, Juji T: **Sequence-based association analysis of HLA class I and II alleles in Japanese supports conservation of common haplotypes.** *Immunogenetics* 1997, **46**(3):199–205.
30. Kawana-Tachikawa A, Tomizawa M, Nunoya J, Shioda T, Kato A, Nakayama EE, Nakamura T, Nagai Y, Iwamoto A: **An efficient and versatile mammalian viral vector system for major histocompatibility complex class I/peptide complexes.** *J Virol* 2002, **76**(23):11982–11988.
31. York-Higgins D, Cheng-Mayer C, Bauer D, Levy JA, Dina D: **Human immunodeficiency virus type 1 cellular host range, replication, and cytopathicity are linked to the envelope region of the viral genome.** *J Virol* 1990, **64**(8):4016–4020.
32. Shioda T, Levy JA, Cheng-Mayer C: **Macrophage and T cell-line tropisms of HIV-1 are determined by specific regions of the envelope gp120 gene.** *Nature* 1991, **349**(6305):167–169.
33. Kuzushima K, Hayashi N, Kimura H, Tsurumi T: **Efficient identification of HLA-A*2402-restricted cytomegalovirus-specific CD8(+) T-cell epitopes by a computer algorithm and an enzyme-linked immunospot assay.** *Blood* 2001, **98**(6):1872–1881.
34. Cole DK, Rizkallah PJ, Gao F, Watson NI, Boulter JM, Bell JI, Sami M, Gao GF, Jakobsen BK: **Crystal structure of HLA-A*2402 complexed with a telomerase peptide.** *Eur J Immunol* 2006, **36**(1):170–179.
35. Otwinowski Z, Minor W: **Processing of X-Ray Diffraction Data Collected in Oscillation Mode Macromolecular Crystallography.** *Methods Enzymol* 1997, **276**:307–326.
36. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM BKE, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS: **Overview of the CCP4 suite and current developments.** *Acta Crystallogr Sect D Biol Crystallgraphy* 2011, **67**:235–242.
37. Vagin A, Teplyakov A: **MOLREP: an Automated Program for Molecular Replacement.** *J Appl Cryst* 1997, **30**:1022–1025.
38. Emsley P, Cowtan K: **Coot: model-building tools for molecular graphics.** *Acta Crystallogr D Biol Crystallogr* 2004, **60**(Pt 12 Pt 1):2126–2132.
39. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC: **Structure validation by Calpha geometry: phi, psi and Cbeta deviation.** *Proteins Struct Funct Genet* 2003, **50**(3):437–450.

doi:10.1186/1742-4690-11-38
Cite this article as: Han *et al.*: Switching and emergence of CTL epitopes in HIV-1 infection. *Retrovirology* 2014 11:38.

APOBEC3D and APOBEC3F Potently Promote HIV-1 Diversification and Evolution in Humanized Mouse Model

Kei Sato[1☯*], Junko S. Takeuchi[1☯], Naoko Misawa[1], Taisuke Izumi[2¤a], Tomoko Kobayashi[1¤b], Yuichi Kimura[1], Shingo Iwami[3], Akifumi Takaori-Kondo[4], Wei-Shau Hu[5], Kazuyuki Aihara[6,7], Mamoru Ito[8], Dong Sung An[9,10,11], Vinay K. Pathak[2], Yoshio Koyanagi[1]

1 Laboratory of Viral Pathogenesis, Institute for Virus Research, Kyoto University, Kyoto, Kyoto, Japan, **2** Viral Mutation Section, HIV Drug Resistance Program, Center for Cancer Research, National Cancer Institute-Frederick, Frederick, Maryland, United States of America, **3** Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka, Fukuoka, Japan, **4** Department of Hematology and Oncology, Graduate School of Medicine, Kyoto University, Kyoto, Kyoto, Japan, **5** Viral Recombination Section, HIV Drug Resistance Program, Center for Cancer Research, National Cancer Institute-Frederick, Frederick, Maryland, United States of America, **6** Institute of Industrial Science, The University of Tokyo, Meguro-ku, Tokyo, Japan, **7** Graduate School of Information Science and Technology, The University of Tokyo, Meguro-ku, Tokyo, Japan, **8** Central Institute for Experimental Animals, Kawasaki, Kanagawa, Japan, **9** Division of Hematology and Oncology, University of California, Los Angeles, Los Angeles, California, United States of America, **10** School of Nursing, University of California, Los Angeles, Los Angeles, California, United States of America, **11** AIDS Institute, University of California, Los Angeles, Los Angeles, California, United States of America

Abstract

Several APOBEC3 proteins, particularly APOBEC3D, APOBEC3F, and APOBEC3G, induce G-to-A hypermutations in HIV-1 genome, and abrogate viral replication in experimental systems, but their relative contributions to controlling viral replication and viral genetic variation *in vivo* have not been elucidated. On the other hand, an HIV-1-encoded protein, Vif, can degrade these APOBEC3 proteins via a ubiquitin/proteasome pathway. Although APOBEC3 proteins have been widely considered as potent restriction factors against HIV-1, it remains unclear which endogenous APOBEC3 protein(s) affect HIV-1 propagation *in vivo*. Here we use a humanized mouse model and HIV-1 with mutations in Vif motifs that are responsible for specific APOBEC3 interactions, DRMR/AAAA (4A) or YRHHY/AAAAA (5A), and demonstrate that endogenous APOBEC3D/F and APOBEC3G exert strong anti-HIV-1 activity *in vivo*. We also show that the growth kinetics of 4A HIV-1 negatively correlated with the expression level of APOBEC3F. Moreover, single genome sequencing analyses of viral RNA in plasma of infected mice reveal that 4A HIV-1 is specifically and significantly diversified. Furthermore, a mutated virus that is capable of using both CCR5 and CXCR4 as entry coreceptor is specifically detected in 4A HIV-1-infected mice. Taken together, our results demonstrate that APOBEC3D/F and APOBEC3G fundamentally work as restriction factors against HIV-1 *in vivo*, but at the same time, that APOBEC3D and APOBEC3F are capable of promoting viral diversification and evolution *in vivo*.

Citation: Sato K, Takeuchi JS, Misawa N, Izumi T, Kobayashi T, et al. (2014) APOBEC3D and APOBEC3F Potently Promote HIV-1 Diversification and Evolution in Humanized Mouse Model. PLoS Pathog 10(10): e1004453. doi:10.1371/journal.ppat.1004453

Editor: Susan R. Ross, University of Pennsylvania School of Medicine, United States of America

Received June 23, 2014; **Accepted** September 5, 2014; **Published** October 16, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported in-part by grants from the following: the Aihara Innovative Mathematical Modelling Project, JSPS through the "Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)," initiated by the Council for Science and Technology Policy of Japan (to KS, SI, and KA); Grants-in-Aid for Scientific Research B24390112 (to YKo) and S22220007 (to MI and YKo) and a Grant-in-Aid for Young Scientists B23790500 (to KS) from the Japan Society for the Promotion of Science (JSPS); a Grant-in-Aid for Scientific Research on Innovative Areas 24115008 (to YKo) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, Research on HIV/AIDS (to YKo) from the Ministry of Health, Labor and Welfare of Japan; Takeda Science Foundation (to KS); Sumitomo Foundation Research Grant (to KS); Senshin Medical Research Foundation (to KS); Imai Memorial Trust for AIDS Research (to KS); JST PRESTO program (to SI); and a UCLA CFAR grant 5P30AI028697 (to DSA); JSPS Research Fellowship for Japanese Biomedical and Behavioral Researchers at NIH (to TI); and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research (to VKP). 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. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors declare that no competing interests exist.

* Email: ksato@virus.kyoto-u.ac.jp

☯ These authors contributed equally to this work.

¤a Current address: Department of Microbiology, Institute of Health Biosciences, The University of Tokushima, Tokushima, Japan
¤b Current address: Laboratory of Animal Health, Department of Animal Science, Faculty of Agriculture, Tokyo University of Agriculture, Atsugi, Kanagawa, Japan

Introduction

Activation-induced cytidine deaminase/apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (AID/APOBEC) superfamily is composed of cellular cytidine deaminases that closely associate with crucial events in vertebrates such as immunity, malignancy, metabolism, and infectious diseases [1,2]. For instance, AID causes somatic hypermutation in B cells resulting in antibody diversification [2], whereas APOBEC1 edits the mRNA of apolipoprotein B and regulates lipid metabolism [3].

Author Summary

Mutation can produce three outcomes in viruses: detrimental, neutral, or beneficial. The first one leads to abrogation of virus replication because of error catastrophe, while the last one lets the virus escape from anti-viral immune system or adapt to the host. Human APOBEC3D, APOBEC3F, and APOBEC3G are cellular cytidine deaminases which cause G-to-A mutations in HIV-1 genome. Here we use a humanized mouse model and demonstrate that endogenous APOBEC3F and APOBEC3G induce G-to-A hypermutation in viral genomes and exert strong anti-HIV-1 activity *in vivo*. We also reveal that endogenous APOBEC3D and/or APOBEC3F induce viral diversification, which can lead to the emergence of a mutated virus that converts its coreceptor usage. Our results suggest that APOBEC3D and APOBEC3F are capable of promoting viral diversification and functional evolution *in vivo*.

The paralogs of human *AID*, *APOBEC1*, and *APOBEC2* genes are encoded in rodents and artiodactyls [4]. On the other hand, although mice encode a sole *Apobec3* gene, primates encode seven paralogs of murine *Apobec3* in their genome, which are designated to *APOBEC3A* to *H*. Given the strong evidence that the duplicated genes have been exposed to selective pressures [5], the seven *APOBEC3* genes have been positively selected [6] and APOBEC3 family proteins play various roles in primates including humans. For instance, APOBEC3A initiates the mutations of foreign DNA (e.g., microbial DNA), which leads to the clearance of bacteria from human cells [7]. In addition, APOBEC3B-mediated mutation closely associates with several human cancers [8,9], particularly breast cancer [10].

APOBEC3G is the most extensively studied APOBEC3 protein in the field of virology and plays a crucial role in the infection and replication of HIV-1, a causative agent of AIDS [11]. APOBEC3G is incorporated into HIV-1 particles and induces G-to-A mutations in the newly synthesized viral DNA, which results in the abrogation of viral replication [4,12]. On the other hand, an HIV-1-encoded protein, viral infectivity factor (Vif), impedes APOBEC3G incorporation into progeny virions by degrading these proteins through the ubiquitin/proteasome-dependent pathway [4,12]. In addition to APOBEC3G, *in vitro* studies using cell culture systems have demonstrated that like APOBEC3G, APOBEC3F and APOBEC3D also potently impair HIV-1 replication [13–15]. However, one study has concluded that APOBEC3F expression levels in T cell lines were not sufficient to inhibit HIV-1 replication [16]. Another study analyzed the replication of HIV-1 Vif mutants that were defective in inducing degradation of APOBEC3G or APOBEC3F in primary $CD4^+$ T cells, and concluded that APOBEC3G exerts a stronger antiviral activity on HIV-1 than APOBEC3F [17]. Thus, the relative impact of different APOBEC3 proteins on HIV-1 replication *in vivo* has not been determined.

Apart from their anti-HIV-1 abilities, certain studies have suggested that APOBEC3-mediated G-to-A mutation can lead to viral evolution and divergence [18–21]. However, it remains unclear how and which endogenous APOBEC3 proteins affect HIV-1 replication, pathogenesis, and diversity *in vivo*.

In order to elucidate the dynamics of HIV-1 infection *in vivo*, we have constructed a humanized mouse model by xenotransplanting human $CD34^+$ hematopoietic stem cells (hHSCs) into an immunodeficient NOD/SCID *Il2rg*$^{-/-}$ (NOG) mouse [22–28]. Our humanized mouse model is able to recapitulate the characteristics of HIV-1 pathogenesis such as the depletion of peripheral $CD4^+$ T cells [22,23,25]. By using this model, we have previously demonstrated that the expression levels of endogenous APOBEC3 genes in human $CD4^+$ T cells of humanized mice were comparable to those of humans and that the combined activity of endogenous APOBEC3 proteins can potently abrogate *vif*-deficient HIV-1 propagation *in vivo* [23]. However, which endogenous APOBEC3 proteins are crucial to the anti-HIV-1 effect *in vivo* is not yet known. In fact, although G-to-A mutations, presumably caused by endogenous APOBEC3 proteins, have been clearly observed in the viral genomes of HIV-1-infected patients, the frequencies of G-to-A mutations seem to vary among individuals and the mutation context is still controversial [21,29–37]. Moreover, because there is a possibility that some endogenous APOBEC3 protein(s) are capable of facilitating viral diversification *in vivo*, it is important to elucidate how endogenous APOBEC3 proteins would affect HIV-1 if we target these molecules for therapy.

In this study, we demonstrate that the propagation of HIV-1 *vif* mutants, which are unable to degrade APOBEC3D/F, APOBEC3G, or both APOBEC3D/F and APOBEC3G, are severely impaired, demonstrating that endogenous APOBEC3D/F and APOBEC3G proteins potently suppress HIV-1 propagation *in vivo*. In addition to the anti-HIV-1 activity of APOBEC3D and APOBEC3F, our results demonstrate that endogenous APOBEC3D and APOBEC3F also potently induced viral diversification. Taken together, our findings show APOBEC3D and APOBEC3F promote HIV-1 diversification *in vivo* and thereby facilitate viral adaptation and evolution.

Results

Strong inhibition of HIV-1 propagation *in vivo* by mutating ^{14}DRMR^{17} and/or ^{40}YRHHY^{44} motifs in Vif

It was demonstrated that ^{14}DRMR^{17} motif in Vif is necessary for the degradation of APOBEC3D and APOBEC3F, while ^{40}YRHHY^{44} motif in Vif is necessary for the degradation of APOBEC3G [38,39]. As shown in Figure 1A, these two motifs were highly conserved in HIV-1 group M. In addition, these motifs are located on the outside regions of Vif protein (Figure 1B) [39]. Moreover, we confirmed that APOBEC3D, APOBEC3F, and APOBEC3G have the ability to decrease *vif*-deficient HIV-1 infectivity *in vitro* (Figure S1).

To confirm the importance of these motifs *in vivo*, we prepared 3 Vif mutants, DRMR/AAAA (4A), YRHHY/AAAAA (5A), and a double mutant (4A5A), based on a CCR5-tropic HIV-1 infectious molecular clone (IMC; strain NLCSFV3) [40]. As previously reported [17], the infectivity of WT, 4A, 5A, and 4A5A HIV-1s were comparable in the absence of APOBEC3s (Figure 1C). On the other hand, the infectivity of 4A HIV-1 was strongly suppressed by APOBEC3D and APOBEC3F but not by APOBEC3G, while that of 5A HIV-1 was decreased by APOBEC3G but not by APOBEC3D and APOBEC3F (Figure 1C). These results indicate that 4A HIV-1 is sensitive to APOBEC3D and APOBEC3F but not to APOBEC3G, while 5A HIV-1 is sensitive to APOBEC3G but not to APOBEC3D and APOBEC3F.

To investigate the anti-HIV-1 activity of each endogenous A3 protein *in vivo*, we inoculated WT, 4A, 5A, and 4A5A HIV-1s into humanized mice. As shown in Figure 2A, the viral loads (VLs) of 4A, 5A, and 4A5A HIV-1s were significantly lower than that of WT HIV-1, and the gradual decrease of peripheral $CD4^+$ T cells was observed only in WT HIV-1-infected mice (Figure 2B). In contrast to *vif*-deleted viruses that did not replicate at all in humanized mouse models [23,41], 4A, 5A and 4A5A HIV-1s

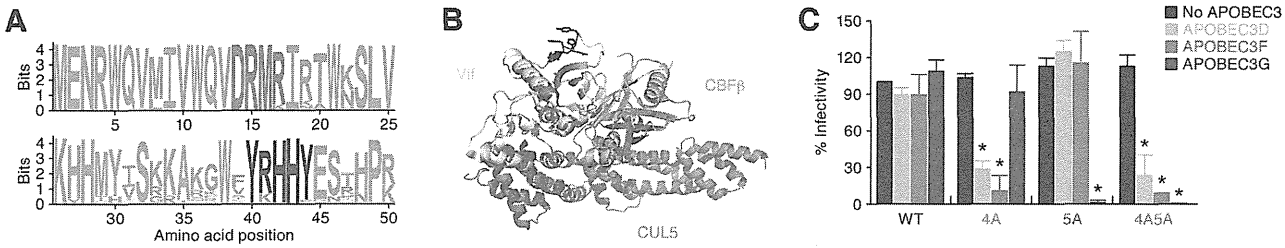


Figure 1. Anti-HIV-1 effect of APOBEC3 proteins *in vitro*. (A) Conservation of ^{14}DRMR^{17} and ^{40}YRHHY^{44} motifs in Vif. The *vif* ORF sequences of HIV-1 group M that are registered in Los Alamos HIV sequence database (n = 7,118) were aligned and analyzed as described in Materials and Methods. (B) Location of DRMR and YRHHY motifs in Vif crystal structure. The 3D structure of Vif was generated on PyMOL v1.6 (http://www.pymol.org/) with the crystal structure of Vif-CBFβ-CUL5-ELOB-ELOC complex (PDB code: 4N9F) [79]. Yellow, cyan, and magenta cartoons respectively represent the main chain of Vif, CBFβ, and CUL5. Red and blue cartoons respectively represent DRMR and YRHHY motifs in Vif. (C) TZM-bl assay. The infectivity of released virions was determined by using TZM-bl cells. The infectivity of each virus is normalized to the value of WT HIV-1 without APOBEC3. The assay was performed in triplicate. *$P<0.05$ versus no APOBEC3 by Student's *t* test. The assay was performed in triplicate. The data represents average with SD.

doi:10.1371/journal.ppat.1004453.g001

exhibited partial viremia (VL at 6 weeks postinfection [wpi]: WT HIV-1, $6.3\times10^5\pm3.0\times10^5$ copies/ml; 4A HIV-1, $1.2\times10^5\pm1.2\times10^5$ copies/ml; 5A HIV-1, $2.9\times10^3\pm0.9\times10^3$ copies/ml; 4A5A HIV-1 $2.3\times10^3\pm0.7\times10^3$ copies/ml). These suggest that the *vif*-mutated viruses used in this study retain some Vif activity, even though they are highly defective. We also inoculated higher doses of viruses into humanized mice; however, despite the higher virus dose, these HIV-1 *vif* mutants did not propagate efficiently *in vivo* (Figures S2A and S2B). These findings strongly suggest that endogenous APOBEC3 proteins, particularly APOBEC3D, APOBEC3F and APOBEC3G, can potently impair HIV-1 propagation *in vivo*.

To quantitatively analyze the magnitude of viral propagation *in vivo*, we evaluated the area under the curve (AUC) of VL (Figure 2C; see also Materials and Methods) and the virus replication rate (Figure 2D; see also Materials and Methods). As shown in Figures 2C and 2D, these two analyses revealed that both the AUC and virus replication rate of WT HIV-1 were significantly higher than those of 4A, 5A, and 4A5A HIV-1s (AUC: WT HIV-1, 10.0 ± 1.8; 4A HIV-1, 2.9 ± 0.7; 5A HIV-1, 1.1 ± 0.3; 4A5A HIV-1, 0.9 ± 0.2. Virus replication rate: WT HIV-1, 15.0 ± 3.0; 4A HIV-1, 4.9 ± 0.8; 5A HIV-1, 2.4 ± 0.4; 4A5A HIV-1, 2.6 ± 0.4). In addition, although the differences in viral load (Figure 2A) and CD4 decline (Figure 2B) between 4A and 5A HIV-1-infected mice were not large, we detected statistically significant differences between 4A and 5A HIV-1-infected mice in AUC (3.8-fold, $P = 0.030$; Figure 2C) and virus replication rate (2.1-fold, $P = 0.0050$; Figure 2D), respectively. Taken together, these findings suggest that endogenous APOBEC3G, APOBEC3F and/or APOBEC3D have the potential to diminish HIV-1 propagation *in vivo* and that the antiviral activity of endogenous APOBEC3G is higher than the combined antiviral activity of APOBEC3D and APOBEC3F.

No reversion of mutations in HIV-1 *vif* mutants

Although the growth of HIV-1 *vif* mutants was generally low, certain mice infected with 4A HIV-1 exhibited moderate levels of viremia (Figure 2A). To assess the possibility that reversion of mutations in *vif* led to the limited spread of HIV-1 *vif* mutants in humanized mice, we analyzed the *vif* mRNA sequences in the spleen of infected mice at 6 wpi. We observed prominent G-to-A mutations in HIV-1 *vif* mutant-infected mice (Figures S3A and S3B). We then asked whether nonsynonymous Vif mutants frequently identified in infected mice (Figure S3C) maintained their ability to degrade APOBEC3 proteins. As shown in Figure 2E, all Vif mutants were expressed at similar levels. However, although some minor variants in 4A HIV-1-infected mice such as K34E (3/320), E88K (5/320), and E134K (9/320) lost their-anti-APOBEC3G activity, all Vif mutants isolated from 4A HIV-1-infected mice were unable to eliminate APOBEC3F, and those from 5A HIV-1-infected mice were unable to eliminate APOBEC3G (Figure 2F). Although I128L mutant (ATA-to-TTA mutation) was predominantly observed in a 4A HIV-1-infected mouse (62 out of the 320 sequences analyzed; Figure S3C), this mutation did not affect its anti-APOBEC3 activity (Figure 2F). These results indicate that the *vif* mutations did not revert the 4A and 5A mutant phenotypes in infected mice, and that the moderate levels of viremia observed in some HIV-1 *vif* mutant-infected mice was not due to the restoration of Vif function.

Negative correlation between the expression level of APOBEC3F and the growth of 4A HIV-1

In the 3 kinds of HIV-1 *vif* mutants, it was noteworthy that the VL in each 4A HIV-1-infected mouse varied between individual mice, while those in 5A and 4A5A HIV-1-infected mice were uniformly low (Figure 3A, left panel). In fact, the coefficient of variance of peak VL, which indicates the extent of distribution, in 4A HIV-1-infected mice was ~2-fold higher than that in WT, 5A, and 4A5A HIV-1-infected mice (Figure 3A, right panel). These findings raised a possibility that the level of viremia in 4A HIV-1-infected mice is correlated with endogenous *APOBEC3* expression levels. To address this possibility, we determined the endogenous expression levels of *APOBEC3D*, *APOBEC3F*, and *APOBEC3G* in the spleen of humanized mice by real-time RT-PCR and standardized them to those of *APOBEC3D* according to the procedure reported previously [42,43]. Since we have previously demonstrated that the spleen is one of the major tissues for HIV-1 replication in our hHSC-transplanted humanized mouse model [25], we assumed that the endogenous expression levels of *APOBEC3* genes in the splenic human mononuclear cells (MNCs) affect the kinetics of 4A HIV-1 growth. As shown in Figure 3B, the expression levels of *APOBEC3F* and *APOBEC3G* were 6.4-fold and 24.7-fold higher than those of *APOBEC3D*, respectively, and the expression level of *APOBEC3G* was 3.9-fold higher than that of *APOBEC3F*.

We then analyzed the endogenous expression level of each *APOBEC3* gene in infected mice. Consistent with a previous study in $CD4^+$ T cell cultures *in vitro* [43], WT HIV-1 infection significantly enhanced the mRNA expression of *APOBEC3D*,

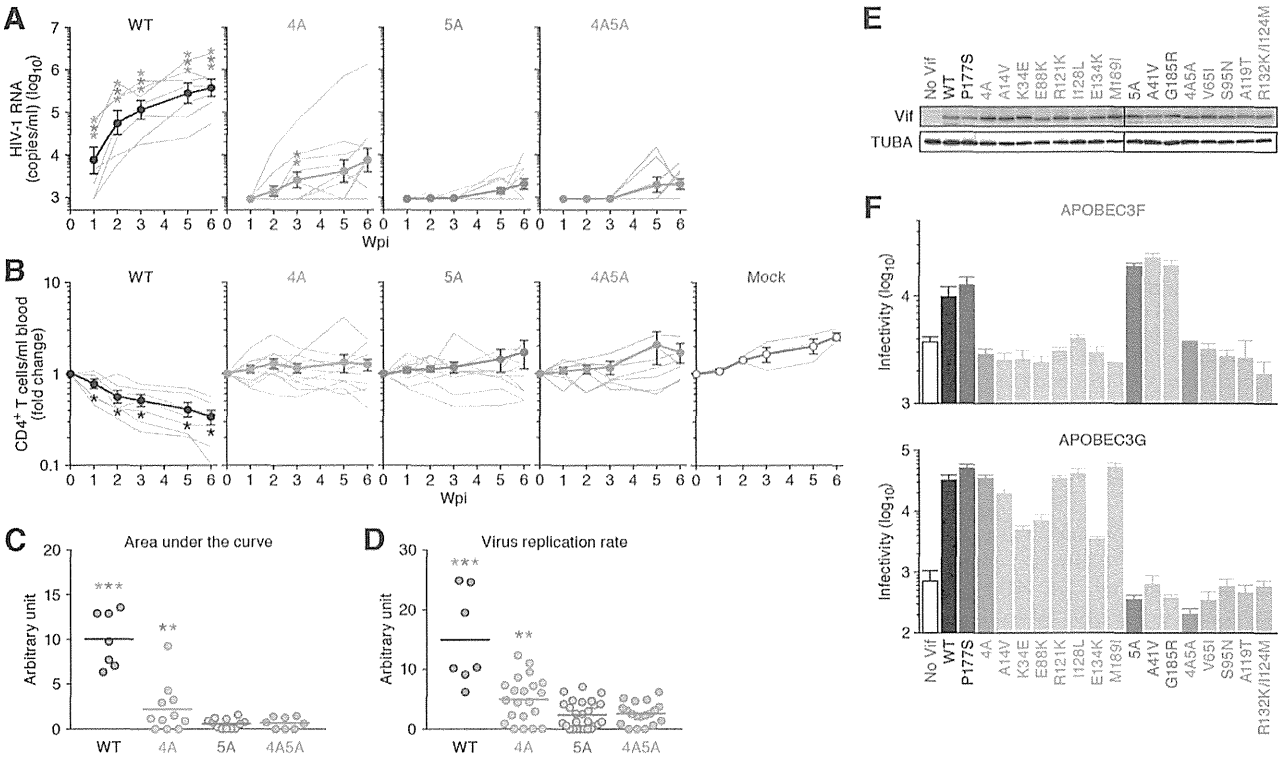


Figure 2. Dynamics of HIV-1 *vif* mutant infection in humanized mice. (A and B) The virus solutions containing 5 ng of p24 antigen (WT HIV-1 [n = 7], 4A HIV-1 [n = 11], 5A HIV-1 [n = 12], and 4A5A HIV-1 [n = 8]) or RPMI 1640 (n = 3; for mock infection) were inoculated into humanized mice, and the amount of viral RNA in plasma (A) and the level of peripheral $CD4^+$ T cells ($CD45^+$ $CD3^+$ $CD4^+$ cells) (B) were analyzed at 0, 1, 2, 3, 5, and 6 wpi. The averages are shown in circles with SEMs, and the values from each mouse are shown by line. X-axes, wpi. In panel A, the detection limit of HIV-1 RNA is 800 copies/ml plasma. (C) Area under the curve (AUC). AUCs of the VL of the mice infected with WT HIV-1 (n = 7), 4A HIV-1 (n = 11), 5A HIV-1 (n = 12), 4A5A HIV-1 (n = 8) were calculated using the trapezoidal rule as described in Materials and Methods. (D) Virus replication rate. Virus replication rates of WT HIV-1 (n = 7), 4A HIV-1 (n = 21), 5A HIV-1 (n = 27), and 4A5A HIV-1 (n = 19) were estimated by using the data of VL and peripheral $CD4^+$ T cell counts as described in Materials and Methods. In panels C and D, horizontal bars represent the averages. Asterisks represent statistically significant differences ($P<0.05$ by Student's *t* test) versus each HIV-1 *vif* mutant (A), between WT HIV-1 and *vif* mutants (C and D), and between infected mice and mock-infected mice (B). In panels A, C, and D, each color of asterisk represents the statistically significant difference against each HIV-1 *vif* mutant-infected mice. (E and F) No Vif reversion in HIV-1 *vif* mutant-infected humanized mice. (E) Western blotting of the Vif mutants frequently observed in infected mice (see also Figure S3). The input of cell lysate was standardized to α-Tubulin (TUBA), and representative results are shown. (F) TZM-bl assay. The expression plasmids of the Vif mutants were cotransfected with pNLCSFV3*Δvif* and either APOBEC3F (*top*) or APOBEC3G (*bottom*) expression plasmids into 293T cells, and the infectivity of released virus was determined by using TZM-bl cells. The assay was performed in triplicate. The data represents average with SD.
doi:10.1371/journal.ppat.1004453.g002

APOBEC3F, and *APOBEC3G* (Figure 3C). Given that HIV-1 infection induces type I interferon (IFN) production in *in vitro* cell cultures [44] and infected individuals during the acute phase [45], taken together with the fact that type I IFNs potently enhance the expression of APOBEC3 genes [42,46], we further evaluated the expression level of *IFNB*, a type I IFN. As shown in Figure 3C, the level of *IFNB* in WT HIV-1-infected mice is also significantly higher than that in mock-infected mice, although the levels of *APOBEC3D*, *APOBEC3F*, *APOBEC3G* and *IFNB* in HIV-1 *vif* mutant infected mice were comparable to mock-infected mice. Moreover, the expression levels of respective *APOBEC3* (Figure S4A) and *IFNB* (Figure S4B) were significantly correlated with each other. These findings suggest that HIV-1 propagation induces type I IFN production resulting in the augmentation of *APOBEC3* expression in humanized mice. In addition to type I IFNs, it has been reported that certain cytokines such as interleukin-2, 7, and 15 [47] and mitogens [43] also potently enhance *APOBEC3* expression. Therefore, the enhancement of *APOBEC3* expression observed in infected humanized mice (Figure 3C) might be a combination effect of type I IFNs and the other factors.

Furthermore, we assessed the relationship between the *APOBEC3* expression level and 4A HIV-1 growth kinetics and found that the VLs in 4A HIV-1-infected mice negatively correlated with the expression level of *APOBEC3F* but not of *APOBEC3D* with statistical significance (Figure 3D; $r = -0.571$, $P = 0.009$). Taken together, these results suggest that the endogenous expression level of *APOBEC3F* determines the growth kinetics of 4A HIV-1 *in vivo*.

Distinct mutation signatures observed in HIV-1 *vif* mutant-infected humanized mice

To analyze the impact of APOBEC3-mediated mutations on the viral genome *in vivo*, we performed semiquantitative differential DNA denaturation PCR (3D-PCR) [48]. In this assay, if G-to-A mutations have accumulated in an amplicon, a PCR product can be detected even at lower denaturation temperatures because the decreased GC content in the amplicon leads to more

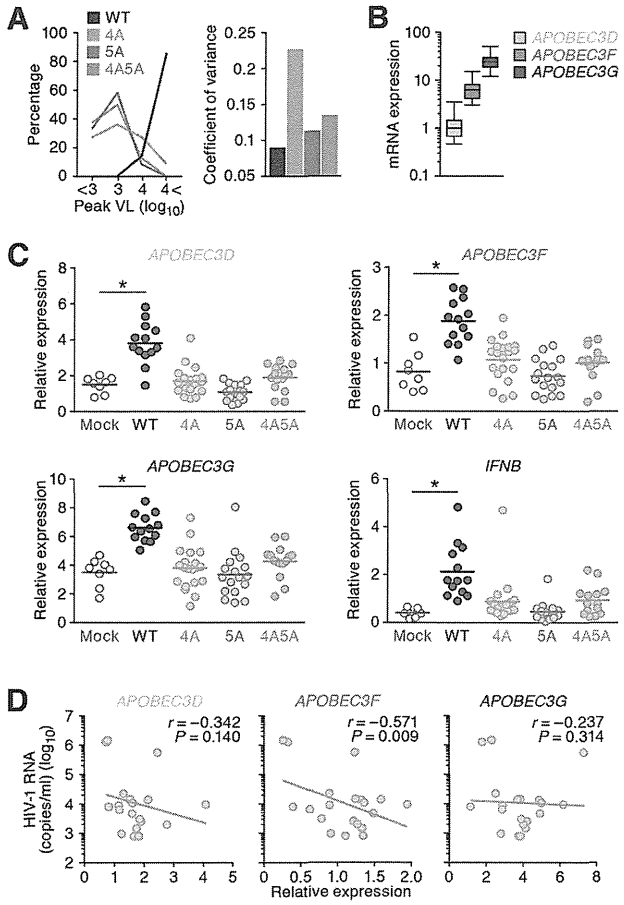


Figure 3. Expression levels of *APOBEC3* and *IFNB* in infected humanized mice. (A) Dispersion of HIV-1 growth efficiency in humanized mice. (*Left*) The peak VLs of WT HIV-1 (n = 7), 4A HIV-1 (n = 11), 5A HIV-1 (n = 12), and 4A5A HIV-1 (n = 8) were classified into 4 degrees (less than 10^3, 10^3–10^4, 10^4–10^5, or more than 10^5), and the distribution is plotted. (*Right*) The coefficient of variance of the peak VLs of WT HIV-1 (n = 7), 4A HIV-1 (n = 11), 5A HIV-1 (n = 12), and 4A5A HIV-1 (n = 8) is shown. (B) Expression levels of *APOBEC3D*, *APOBEC3F*, and *APOBEC3G* in the splenic human $CD4^+$ T cells of humanized mice (n = 73) were analyzed by real-time RT-PCR. The values were standardized as previously described [43], and the level of *APOBEC3D* is set to 1 to facilitate comparison. (C) Expression levels of *APOBEC3D*, *APOBEC3F*, *APOBEC3G*, and *IFNB* in the splenic human $CD4^+$ T cells of infected mice (WT, n = 13; 4A, n = 20; 5A, n = 17; and 4A5A, n = 15) and mock-infected mice (n = 8) at 6 wpi were analyzed by real-time RT-PCR. Horizontal bars represent the averages. Asterisks represent statistically significant difference ($P<0.05$ by Student's *t* test) between infected mice and mock-infected mice. (D) Negative correlation between VL and *APOBEC3* expression in 4A HIV-1-infected humanized mice. The mRNA expression levels of *APOBEC3D* (*left*), *APOBEC3F* (*middle*), and *APOBEC3G* (*right*) in the splenic human $CD4^+$ T cells (x-axes) and the VL at 6 wpi (y-axis) of 4A HIV-1-infected mice (n = 20) are shown. The lines represent exponential approximation. Pearson correlation coefficient (*r*) was adopted to determine statistically significant correlation between each value.

doi:10.1371/journal.ppat.1004453.g003

efficient denaturation at lower temperatures [48,49]. As shown in Figure 4A, the 3D-PCR products were detected at relatively lower denaturation temperatures in HIV-1 *vif* mutants but not of WT HIV-1, suggesting that the proviral genomes of HIV-1 *vif* mutants suffered from APOBEC3-mediated G-to-A hypermutation *in vivo*.

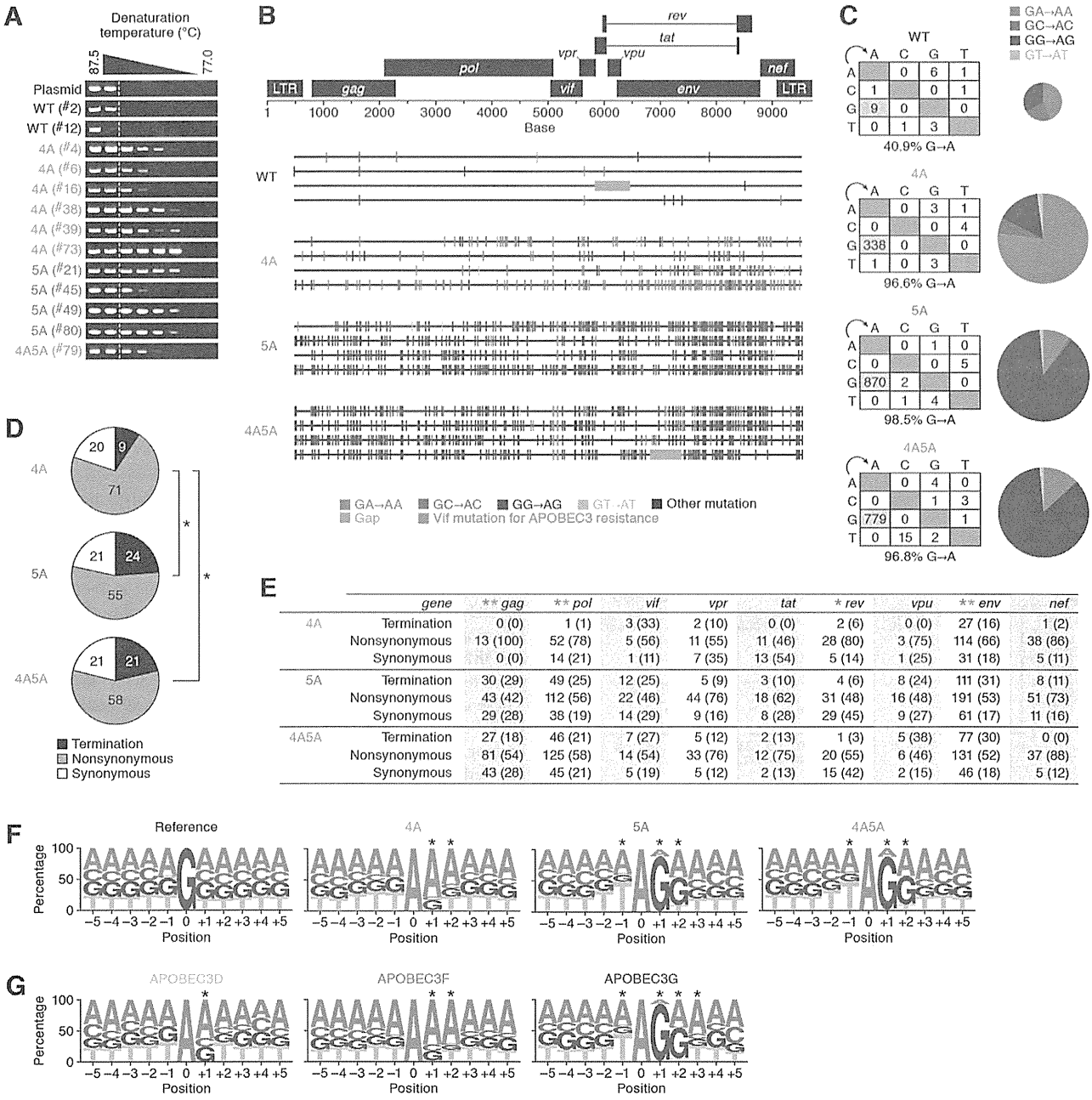
It is known that APOBEC3G predominantly generates GG-to-AG mutations, while APOBEC3D and APOBEC3F predominantly generate GA-to-AA mutations [14]. We assessed the sequence of full-length proviral DNA in the spleen of infected mice at 6 wpi, and found that GA-to-AA hypermutation was frequently observed in 4A HIV-1, while GG-to-AG hypermutation was readily observed in 5A and 4A5A HIV-1s (Figures 4B and 4C).

We then assessed the effect of G-to-A mutation detected in the proviral DNA of infected mice. As shown in Figure 4D, the results revealed that the percentage of termination codon mutations in 4A HIV-1 (9.3%) was significantly lower than those in 5A HIV-1 (23.8%) and 4A5A HIV-1 (21.3%) (4A HIV-1 versus 5A HIV-1, $P = 0.62\times10^{-9}$; 4A HIV-1 versus 4A5A HIV-1, $P = 0.41\times10^{-6}$ by Chi-square test for independence). Similar results were observed in the longer viral genes such as *gag*, *pol*, and *env* (Figure 4E). These results strongly suggest that APOBEC3G efficiently generates termination codons compared to APOBEC3D and APOBEC3F.

To investigate the trend of G-to-A mutation sites in depth, we verified the nucleotides positioned between −5 to +5 from the G-to-A mutation sites in the proviral DNA. Comparing the observed mutations to the expected random G-to-A mutations (shown as "reference" in Figure 4F), statistical analyses revealed that the mutation signature of 4A HIV-1-infected mice is GAA-to-AAA, while those of 5A and 4A5A HIV-1-infected mice were TGGG-to-TAGG (Figures 4F and S5A). Moreover, *in vitro* single-round infection assays revealed that the mutation signatures of APOBEC3D, APOBEC3F, and APOBEC3G were GA-to-AA, GAA-to-AAA, and TGGG-to-TAGG, respectively (Figure 4G). These results indicate that the mutation signature observed in 4A HIV-1-infected was statistically similar to those of APOBEC3D and/or APOBEC3F, and that those in 5A and 4A5A HIV-1-infected mice were statistically similar to that of APOBEC3G (Figure S5B).

Diversification of 4A HIV-1 *in vivo*

Our findings in both *in vivo* (Figure 4F) and *in vitro* (Figure 4G) demonstrated that APOBEC3G prefers to target TGGG as substrate. Importantly, TGG and TAG are the codons encoding Tryptophan and termination codon, respectively, suggesting that APOBEC3G can readily cause lethal mutations (i.e., TGG-to-TAG termination mutations). On the other hand, APOBEC3F and APOBEC3D generated GAA-to-AAA and GA-to-AA mutations, respectively (Figures 4F and 4G), which do not generate termination codons and thus cause lethal mutations less frequently. These findings raised a hypothesis that APOBEC3G directly causes lethal mutations, while APOBEC3F and APOBEC3D induce the accumulation of nonsynonymous mutations in the viral genome. To address this hypothesis, single genome sequencing (SGS) assays [50] were performed using viral RNA isolated from the plasma of infected mice at 6 wpi. Since G-to-A mutations were frequently observed in the proximal upstream region of the 3′ polypurine tract (positioned at 9056–9071; Figure S6), which was consistent with previous reports [49,51], we focused on the *env* open reading frame (ORF) sequence. As shown in Figure 5A (the raw data is shown in Figure S7), SGS assay revealed that G-to-A mutations were frequently observed in the viral RNA genomes of 4A HIV-1-infected mice but not of WT, 5A, and 4A5A HIV-1-infected mice. In addition, in the 91 *env* amplicons of 4A HIV-1-infected mice, 37 analyzed amplicons harbored more than 10 G-to-A mutations, and 15 analyzed amplicons harbored more than 10 GA-to-AA mutations, respectively (Figure S8). On the other hand, the amplicons harboring G-to-A hypermutations were



	gene	** gag	** pol	vif	vpr	tat	* rev	vpu	** env	nef
4A	Termination	0 (0)	1 (1)	3 (33)	2 (10)	0 (0)	2 (6)	0 (0)	27 (16)	1 (2)
	Nonsynonymous	13 (100)	52 (78)	5 (56)	11 (55)	11 (46)	28 (80)	3 (75)	114 (66)	38 (86)
	Synonymous	0 (0)	14 (21)	1 (11)	7 (35)	13 (54)	5 (14)	1 (25)	31 (18)	5 (11)
5A	Termination	30 (29)	49 (25)	12 (25)	5 (9)	3 (10)	4 (6)	8 (24)	111 (31)	8 (11)
	Nonsynonymous	43 (42)	112 (56)	22 (46)	44 (76)	18 (62)	31 (48)	16 (48)	191 (53)	51 (73)
	Synonymous	29 (28)	38 (19)	14 (29)	9 (16)	8 (28)	29 (45)	9 (27)	61 (17)	11 (16)
4A5A	Termination	27 (18)	46 (21)	7 (27)	5 (12)	2 (13)	1 (3)	5 (38)	77 (30)	0 (0)
	Nonsynonymous	81 (54)	125 (58)	14 (54)	33 (76)	12 (75)	20 (55)	6 (46)	131 (52)	37 (88)
	Synonymous	43 (28)	45 (21)	5 (19)	5 (12)	2 (13)	15 (42)	2 (15)	46 (18)	5 (12)

Figure 4. G-to-A hypermutation in the proviral DNA of infected humanized mice. (A) Semiquantitative 3D-PCR. Representative results are shown. Mouse IDs are shown in parentheses (correspond to those in Table S2). The dotted line indicates the lowest denaturation temperature (86.8°C) at which the PCR product is amplified from WT HIV-1-infected mice (mouse ID #2). (B and C) Full-length proviral DNA were cloned and sequenced as described in Materials and Methods. Representative results (B), mutation matrix (C, *left*), and pie chart of G-to-A mutation (C, *right*) are respectively shown. The diameters of pie charts represent the percentage of G-to-A mutations in total mutations. (D and E) Effect of G-to-A mutation in proviral DNA. (D) Pie chart of the effect of G-to-A mutation in full-length proviral DNA. The numbers in pie chart represent the percentage of termination, nonsynonymous, and synonymous mutations in G-to-A mutations, respectively. (E) Summary of the effect of G-to-A mutation in proviral DNA. The numbers and the percentages (in parentheses) of termination, nonsynonymous, and synonymous G-to-A mutations in each viral gene are summarized. Asterisks represent statistically significant differences ($P<0.01$ by Chi-square test for independence). In panel E, each color of asterisk represents the statistically significant difference between 4A HIV-1 and each HIV-1 *vif* mutant. (F and G) G-to-A mutation sites in the proviral DNA of infected mice (F) and *in vitro* infection assay (G) were respectively classified according to the nucleotides positioned between −5 to +5 from the detected G-to-A mutation sites (position 0). The results were respectively compared to that expected if G-to-A mutations occurred randomly occurred (F, 'reference'). *$P<0.001$ between the obtained and the expected results in each position by Chi-square test for independence. See also Figure S5.
doi:10.1371/journal.ppat.1004453.g004

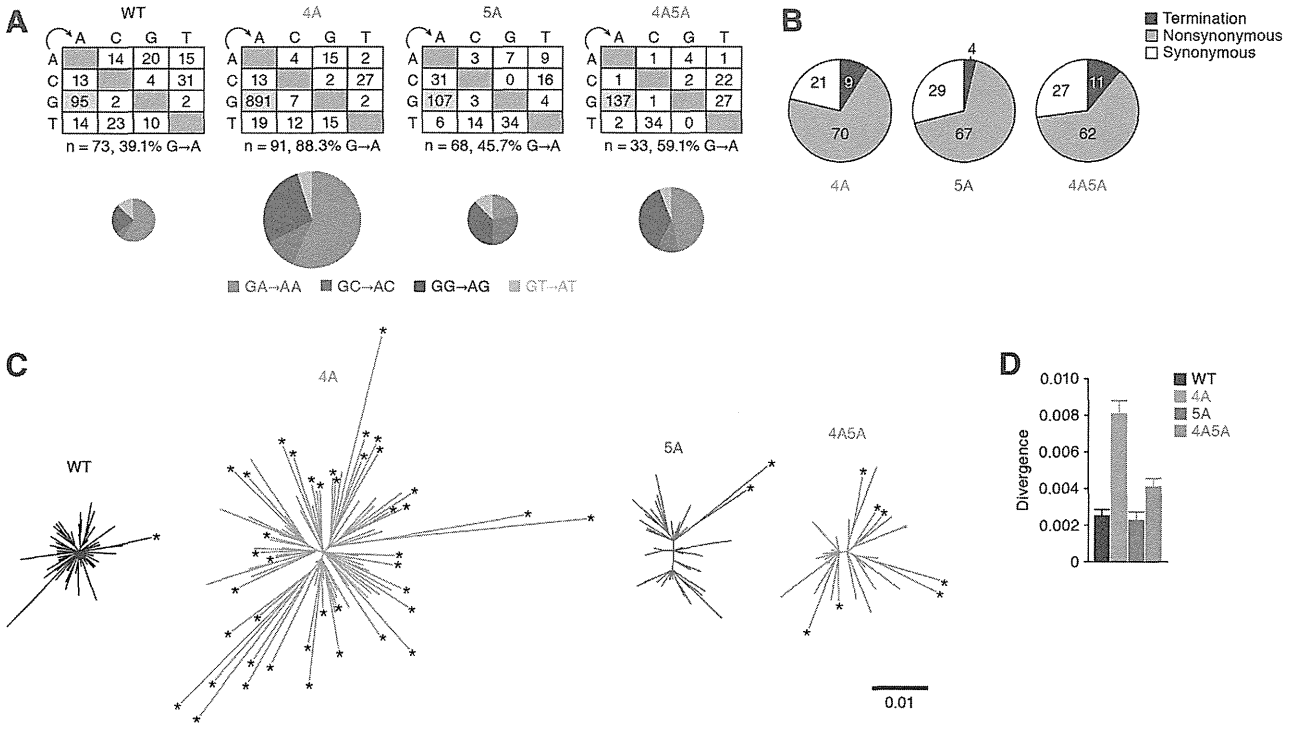


Figure 5. Diversification and functional evolution of 4A HIV-1 *in vivo*. The *env* ORFs (6221–8782, 2,562 bases) of viral RNA in the plasma of infected mice (WT, n = 73 from 2 mice; 4A, n = 91 from 3 mice; 5A, n = 68 from 2 mice; and 4A5A, n = 33 from 1 mouse) were sequenced by SGS assay. Raw data are shown in Figure S7. (A) The mutation matrix (*top*) and the pie chart of G-to-A mutation (*bottom*) are shown. In the bottom panel, the diameters of pie charts represent the percentage of G-to-A mutations in total mutations. (B) Effect of G-to-A mutation in *env* ORF of viral RNA in plasma. Pie chart of the effect of G-to-A mutation in *env* ORF is shown. The numbers in pie chart represent the percentage of termination, nonsynonymous, and synonymous mutations in G-to-A mutations, respectively. (C and D) Divergence of viral RNA sequence. Phylogenic trees (C) and genetic diversity (D) of *env* ORF sequences in the plasma of infected mice are shown. In panel C, the scale bar indicates the number of substitutions per site. The amplicons harboring statistically significant levels of G-to-A mutations ($P<0.05$ by Fisher's exact test using Hypermut 2.0) are indicated by asterisks.
doi:10.1371/journal.ppat.1004453.g005

rarely detected in WT, 5A, and 4A5A HIV-1-infected mice (Figure S8). Moreover, although termination mutations were prominently detected in the proviral DNA of 5A and 4A5A HIV-1-infected mice (Figures 4D and 4E), the percentages of termination mutation in the viral RNA in plasma of 5A and 4A5A HIV-1-infected mice were comparable to that of 4A HIV-1-infected mice (Figure 5B; 4A HIV-1 versus 5A HIV-1, $P = 0.06$; 4A HIV-1 versus 4A5A HIV-1, $P = 0.19$ by Chi-square test for independence). These findings strongly suggest that APOBEC3G-mediated G-to-A mutations frequently result in lethal mutations.

Interestingly, the phylogenic trees displayed that the *env* sequences in 4A HIV-1-infected mice were highly divergent and harbored significant levels of G-to-A mutations (Figure 5C). Furthermore, the analyses on genetic distance directly demonstrated that the *env* RNA sequences of 4A HIV-1-infected mice were highly divergent when compared to those of WT, 5A, and 4A5A HIV-1-infected mice (Figure 5D). Taken together, these findings provide strong evidence that APOBEC3F and APOBEC3D have the potential to restrict HIV-1 propagation, but at the same time, can also augment the emergence of quasispecies through sub-lethal G-to-A mutations *in vivo*.

Emergence of CCR5/CXCR4 dual-tropic HIV-1 in 4A HIV-1-infected mice

As shown in Figure 6A, mutations were detected in both conserved and variable regions of *env*. Previous studies have demonstrated that the variable region 3 (V3) of *env*, particularly the residues positioned at 11 and 25 in the V3, determines the CCR5 or CXCR4 coreceptor usage for HIV-1 entry [52,53]. Since we detected the diversified *env* sequences particularly in 4A HIV-1-infected mice (Figure 5C), we hypothesized the emergence of viruses that can use CXCR4 as the coreceptor in 4A HIV-1-infected mice. To address this possibility, we screened putative CXCR4-tropic HIV-1 by using a geno2pheno tool, which predicts the coreceptor usage based on nucleotide sequence [54], and found that the frequency of putative CXCR4-tropic HIV-1 in 4A HIV-1-infected mice was significantly higher than those in mice infected with WT, 5A, and 4A5A HIV-1s (Figure 6B, *left*). The detected putative CXCR4-tropic viruses were a N7S mutant from a WT HIV-1-infected mouse, a G24R mutant from a 4A HIV-1-infected mouse, and five E25K mutants from three 4A HIV-1-infected and one 4A5A HIV-1-infected mice (Figure 6B, *right*). It was particularly noteworthy that the E25K mutant detected was due to GAA-to-AAA mutation, which is the mutation signature mediated by APOBEC3F and APOBEC3D. To functionally evaluate whether these mutants can use CXCR4 as the coreceptor, we prepared the mutated virus based on NLCSFV3, which exclusively use CCR5 as the coreceptor. As shown in Figure 6C, we directly demonstrated that the infectivity of E25K mutant in $CXCR4^+$ MaRBLE cells was 2.5-fold higher than that of parental NLCSFV3 with a statistical significance ($P = 0.0073$). Taken together, these findings strongly suggest that the G-to-A

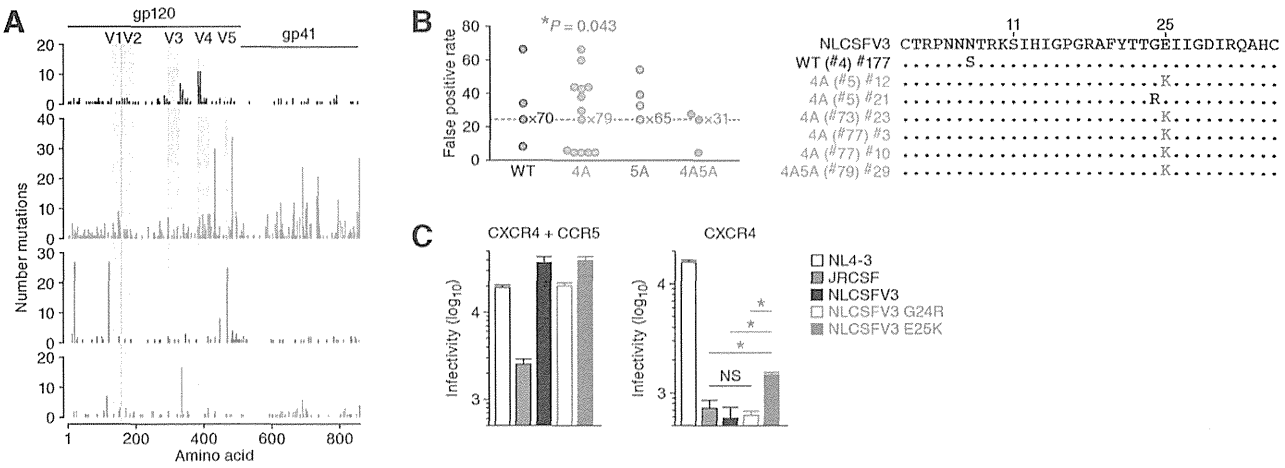


Figure 6. Functional evolution of 4A HIV-1 *in vivo*. (A) Mutations in *env* ORFs. (B) Estimation of coreceptor usage. Putative coreceptor usage was determined by using a geno2pheno coreceptor algorithm [54] as described in Materials and Methods. Mouse IDs are shown in parentheses (correspond to those in Table S2). The frequency of putative CXCR4-tropic HIV-1 in 4A HIV-1-infected mice was significantly higher than those in the mice infected with the other viruses ($P = 0.043$ by Chi-square test for independence). (C) Functional evaluation of coreceptor usage. Viral infectivity was measured by MaRBLE assay using R5-MaRBLE cells (*left*) and X4-MaRBLE cells (*right*). The data represents average with SD. The assay was performed in triplicate. Asterisks represent statistically significant differences ($P<0.05$ by Student's *t* test). NS, no statistical significance.
doi:10.1371/journal.ppat.1004453.g006

mutation mediated by APOBEC3F and APOBEC3D can contribute to the conversion of viral coreceptor usage from CCR5 to CXCR4.

Discussion

Previous studies including ours have demonstrated that endogenous APOBEC3 proteins have robust potential to diminish HIV-1 replication in humanized mouse models [23,41]. Furthermore, Krisko et al. have demonstrated that greater than 80% of G-to-A mutations in their *in vivo* experiments were in the context of GG-to-AG mutations, suggesting that endogenous APOBEC3G is the dominant restricting factor *in vivo* [41]. However, there are no reports that directly evaluate and compare the sole effects of endogenous APOBEC3G and/or APOBEC3D/F on HIV-1 replication *in vivo*. In addition, these papers [23,41] did not explore the possibility that endogenous APOBEC3 protein(s) may contribute to viral diversification. In our present study, we directly examined these two issues by using 3 kinds of HIV-1 *vif* mutants and a humanized mouse model. We demonstrated that endogenous APOBEC3G and APOBEC3D/F are intrinsic restriction factors against HIV-1. Moreover, we observed that endogenous APOBEC3D and APOBEC3F are capable of enhancing viral diversification *in vivo*.

We found that the propagation of HIV-1 *vif* mutants, particularly 5A and 4A5A, was severely suppressed even at high doses (Figures 2A–2D and S2). Consistent with previous reports [42,43], endogenous *APOBEC3G* was highly expressed in the splenic human $CD4^+$ T cells of humanized mice when compared to *APOBEC3F* and *APOBEC3D* (Figure 3B). In addition, the proviral DNA of 5A and 4A5A HIV-1-infected mice exhibited TGGG-to-TAGG hypermutations (Figure 4F), and APOBEC3G preferentially targeted to the tetranucleotide TGGG as substrate (Figure 4G), which readily results in termination codon mutations (Figures 4D and 4E). These findings indicate that endogenous APOBEC3G is an intrinsic factor that severely restricts HIV-1 propagation *in vivo*.

It was notable that the proviral DNA (Figures 4B–4F) and the *vif* ORF in the spleens of 4A5A HIV-1-infected mice (Figure S3) exhibited the signature of APOBEC3G-mediated mutations. In this regard, an *in silico* study has been recently reported that APOBEC3G and APOBEC3F rarely co-mutate the same viral genome in infected individuals [55]. Because the expression level of *APOBEC3G* was higher than those of *APOBEC3F* and *APOBEC3D* (Figure 3B), our findings suggest that APOBEC3G more predominantly affects HIV-1 replication *in vivo* than APOBEC3F and APOBEC3D.

When compared to APOBEC3G, the potential role of APOBEC3D and APOBEC3F in inhibition of viral replication has been controversial. Refsland and colleagues have recently demonstrated the anti-HIV-1 ability of APOBEC3D and APOBEC3F endogenously expressed in a human $CD4^+$ T cell line called CEM2n cells [15]. On the other hand, certain previous studies using human PBMC *in vitro* cultures have suggested that endogenously expressed APOBEC3F moderately restricts [56] or does not restrict [16] *vif*-deficient HIV-1 replication. In this regard, it should be noted that *in vitro* culture conditions use human $CD4^+$ T cell lines and/or human PBMCs artificially activated with mitogens such as phytohemagglutinin, which may not exactly mimic *in vivo* conditions, and therefore, may not reproduce the expression levels of APOBEC3D and APOBEC3F *in vivo*. Thus, it was important to carry out the *in vivo* experiments, which now firmly establish that APOBEC3D/F do indeed exert a substantial anti-viral effect on HIV-1 replication (Figures 2A–2D).

Although we demonstrated that the growth kinetics of 4A HIV-1 was significantly impaired compared to WT HIV-1 (Figures 2A–2D), the kinetics of 4A HIV-1 varied in each mouse (Figure 3A) and were significantly higher than those of 5A and 4A5A HIV-1 (Figures 2C and 2D). Also, the growth kinetics of 4A HIV-1 significantly and negatively correlated to the expression level of *APOBEC3F* but not *APOBEC3D* (Figure 3D), suggesting that endogenous APOBEC3F more critically modulates 4A HIV-1 replication *in vivo* than APOBEC3D. In fact, endogenous expression level of *APOBEC3F* was higher than that of *APOBEC3D* (Figure 3B). Moreover, anti-HIV-1 activity of APOBEC3F was higher than that of APOBEC3D in *in vitro*

transfection experiments (Figures 1C and S1), which are consistent with previous reports [13,17,48]. Therefore, these results suggest that the growth kinetics of 4A HIV-1 is predominantly impaired by APOBEC3F rather than APOBEC3D.

It is known that certain APOBEC3 proteins can impair HIV-1 replication by inhibiting viral reverse transcription (RT) independently of their deaminase activities [57–60]. In this regard, based on an experimental-mathematical approach, we have recently demonstrated that APOBEC3G restricts HIV-1 replication almost completely in a deaminase activity-dependent manner, while APOBEC3F impairs viral replication with the combination of G-to-A mutations and inhibition of viral RT [61]. In addition, although a deaminase-defective APOBEC3G mutant (E259Q) severely lost its anti-viral effect by 173-fold, the anti-viral effect of the deaminase-defective APOBEC3D (E264Q) and APOBEC3F (E251Q) differed only 2–3-fold compared to the WT proteins (Figure S1). These findings suggest that the anti-viral effect of APOBEC3D and APOBEC3F may be partially attributed to their deaminase-independent properties. On the other hand, Albin et al. have recently reported that *vif*-deficient HIV-1 can overcome the anti-viral effect of deaminase-defective APOBEC3F in a spreading infection experiment using T cell lines, and that APOBEC3F's deaminase activity is crucial for long-term restriction of *vif*-deficient HIV-1 replication [62]. Moreover, Mbisa et al. previously reported that virion-incorporated APOBEC3F and APOBEC3G potently inhibit HIV-1 integration [63]. Thus, these findings indicate that APOBEC3 proteins potently suppress HIV-1 replication by at least 3 different modes: (i) G-to-A mutation; (ii) inhibition of viral RT; and (iii) inhibition of viral integration; moreover, the magnitude of each mode of inhibition may be different for specific APOBEC3 proteins. Because APOBEC3's anti-viral modes are complex and intertwined, it would be technically impossible to quantitatively elucidate this under *in vivo* conditions. However, when compared to the mutation signature of APOBEC3G (T<u>G</u>GG-to-T<u>A</u>GG), APOBEC3D and APOBEC3F respectively preferred the dinucleotide (<u>G</u>A) and trinucleotide (<u>G</u>AA), which rarely led to stop codon mutations (Figure 4G) [61]. Although the extent of deaminase-dependent anti-HIV-1 activity of APOBEC3 proteins *in vivo* remains undetermined, our results suggest that endogenous APOBEC3D and APOBEC3F may inhibit HIV-1 replication *in vivo* in a manner that is less dependent on their deaminase activity than APOBEC3G.

Separate from the anti-HIV-1 ability of APOBEC3 proteins, some papers have suggested that the mutations generated by APOBEC3 proteins, particularly APOBEC3G, can promote viral evolution [18,19,21]. In this regard, it was particularly noteworthy that the viral RNA sequences in the plasma of 4A HIV-1-infected mice were highly diversified when compared to those of WT, 5A, and 4A5A HIV-1-infected mice (Figures 5C and 5D). These findings suggest that the G-to-A mutations mediated by APOBEC3D and APOBEC3F, but not by APOBEC3G, can increase the genetic diversity of viral populations. In fact, here we directly showed the emergence of CCR5/CXCR4 dual-tropic HIV-1 most exclusively in 4A HIV-1-infected mice (4 out of the 91 amplicons analyzed; Figure 6C, *right*), and the 4 E25K amplicons detected had intact ORFs (i.e., no termination mutations in the amplicon). More importantly, the E25K mutant in the V3 region of *env*, a coreceptor-switched HIV-1, was generated by a <u>G</u>AA-to-<u>A</u>AA mutation, strongly suggesting that this mutation may be caused by APOBEC3F, and also possibly APOBEC3D. Regarding viral coreceptor usage, it is well known that the charge of two specific amino acids in the V3 region of HIV-1 Env, positioned at 11 and 25, strongly influence the coreceptor usage [52,53]. Our findings strongly suggest that one of the two crucial mutations, E25K, needed for conversion of CCR5 to CXCR4 usage, is facilitated by APOBEC3D/F. This makes it more likely that coreceptor conversion will occur as a result of a random RT error leading to a substitution at the position 11 in genomes that have the APOBEC3D/F-associated mutation.

In addition to the conversion of coreceptor usage (Figure 6C), we found that the sites preferred by APOBEC3D and APOBEC3F may potentially lead to the resistance to anti-HIV-1 drugs (Table S1). Furthermore, our results suggest that 4A HIV-1 can propagate *in vivo* when APOBEC3F expression level was relatively low (Figure 3D). Our data further suggest that sub-lethal G-to-A mutations caused by endogenous APOBEC3D and APOBEC3F, which are expressed at a lower level, rather than APOBEC3G, can lead to diversification of HIV-1 genomes leading to increased viral variation and evolutionary potential. Although our viruses used in this study produce defective Vifs, we believe it reflects the natural infection because Simon et al. have shown that defective *vif*s are often seen during natural infection in patients [64]. Thus, the types of mutations and diversification we observed in this study would be quantitatively higher, but similar to the diversification that occurs during natural infection, as a result of the emergence of *vif*-mutated viruses.

In conclusion, here we demonstrated that endogenous APOBEC3G is the *bona fide* anti-HIV-1 restriction factor even *in vivo*. On the other hand, we also provide strong evidence indicating that endogenous APOBEC3D and APOBEC3F suppress viral replication *in vivo*, while these proteins potently induce viral evolution. These findings suggest that the impairment of Vif-APOBEC3G interaction can be a novel target for anti-HIV-1 drugs, while the restoration of deaminase activity of APOBEC3D and APOBEC3F by inhibiting Vif-mediated degradation may potentially lead to the enhancement of viral diversification. As shown in Figure 1B, both DRMR and YRHHY motifs are exposed on the surface of Vif protein [39]. Therefore, it may be possible to design compounds that target the YRHHY motif and specifically block Vif-APOBEC3G interaction, which may be ideal candidates for development of novel anti-HIV-1 drugs.

Materials and Methods

Ethics statement

All procedures including animal studies were conducted following the guidelines for the Care and Use of Laboratory Animals of the Ministry of Education, Culture, Sports, Science and Technology, Japan. The authors received approval from the Institutional Animal Care and Use Committees (IACUC)/ethics committee of Kyoto University institutional review board (protocol number D13-25). All protocols involving human subjects were reviewed and approved by the Kyoto University institutional review board. Informed written consent from human subjects was obtained in this study.

Humanized mice

NOG mice [65] were obtained from the Central Institute for Experimental Animals (Kawasaki, Kanagawa, Japan). The mice were maintained under specific-pathogen-free conditions and were handled in accordance with the regulation of IACUC/ethics committee of Kyoto University. Human $CD34^+$ hematopoietic stem cells were isolated from human fetal liver as previously described [66]. The humanized mouse (NOG-hCD34 mouse) was constructed as previously described [22–27]. Briefly, 82 newborn (aged 0 to 2 days) NOG mice from 19 litters were irradiated with X-ray (10 cGy per mouse) by an RX-650 X-ray cabinet system

(Faxitron X-ray Corporation) and were then intrahepatically injected with the obtained human fetal liver-derived $CD34^+$ cells (8×10^4 to 17×10^4 cells). A list of the humanized mice used in this study is summarized in Table S2.

Cell culture

293T cells and TZM-bl cells (obtained through the NIH AIDS Research and Reference Reagent program) [67] were maintained in DMEM containing 10% fetal calf serum (FCS) and antibiotics. X4-MaRBLE and R5-MaRBLE cells (kindly provided by Dr. Wataru Sugiura) [68] were maintained in RPMI 1640 containing 10% FCS and antibiotics. For X4-MaRBLE cells, 250 μg/ml Geneticin and 0.1 μg/ml Puromycin were added to the culture medium. For R5-MaRBLE cells, 150 μg/ml Hygromycin B, 250 μg/ml Geneticin, and 0.1 μg/ml Puromycin were added to the culture medium.

Virus preparation and infection

IMCs of CCR5-tropic HIV-1 (strain NLCSFV3) [40] and its derivatives were constructed based on pNLCSFV3 [40]. To construct pNLCSFV3-DRMR/AAAA (pNLCSFV3-4A), pNLCSFV3-YRHHY/AAAAA (pNLCSFV3-5A), and pNLCSFV3*Δvif*, the *Age*I-*Eco*RI fragments of pNL4-3-based these mutants [17,38,69] and pNL4-3*Δvif* [70] were subcloned into the *Age*I-*Eco*RI site of pNLCSFV3 [40]. To construct pNLCSFV3-4A5A, the 5A mutation was inserted into pNLCSFV3-4A as previously described [69]. The sequences of these constructed plasmids were confirmed by sequencing PCR. To prepare the virus solutions for the experiments using humanized mice, 30 μg of pNLCSFV3 or its derivatives (pNLCSFV3-4A, pNLCSFV3-5A, or pNLCSFV3-4A5A) was transfected into 293T cells by the calcium-phosphate method as previously described [23,25]. After 48 h posttransfection, the culture supernatant was harvested, centrifuged, and then filtrated through a 0.45-μm filter (Millipore) to produce virus solution. The amount of virus particles was quantified by using an HIV-1 p24 antigen ELISA kit (Zeptometrix), and 50% infectious dose (ID_{50}) was measured by Reed-Meunch's method as previously described [25]. Virus solutions containing 5 ng (Figure 2), 50 ng and 500 ng (Figure S2) of p24 antigen (equivalent to 1,500, 15,000, and 150,000 ID_{50}, respectively) were intraperitoneally inoculated into NOG-hCD34 mice. RPMI 1640 was used for mock infection.

Peripheral blood collection, mononuclear cell isolation, and quantification of HIV-1 RNA in plasma

Peripheral blood and plasma were collected at 0, 1, 2, 3, 5, and 6 wpi as previously described [22,23,25,26]. The mice were sacrificed at 6 wpi with anesthesia, and the spleen was crushed, rubbed, and suspended as previously described [22,23,25,26]. To obtain splenic human MNCs, the splenic cell suspension was separated by using Ficoll-Paque (Pharmacia) as previously described [22,23,25,26]. The amount of HIV-1 RNA in 50 μl plasma was quantified by Bio Medical Laboratories, Inc. (the detection limit of HIV-1 RNA is 800 copies/ml).

Flow cytometry and hematometry

Flow cytometry was performed with a FACS Canto II (BD biosciences) as previously described [22,23,25,26], and the obtained data were analyzed with Cell Quest software (BD biosciences) and FlowJo software (Tree Star, Inc.). For flow cytometry analysis, anti-CD45-PE (HI30; Biolegend), anti-CD3-APC-Cy7 (HIT3a; Biolegend), and anti-CD4-APC (RPA-T4; Biolegend) antibodies were used. Hematometry was performed with a Celltac α MEK-6450 (Nihon kohden, Co.) as previously described [23,25,26].

Transfection, western blotting, TZM-bl assay, and MaRBLE assay

In vitro transfection experiments were performed by using Lipofectamine 2000 (Life technologies) according to the manufacture's protocol. After 48 h posttransfection, the culture supernatant was harvested, centrifuged, and then filtrated through a 0.45-μm filter (Millipore) to produce virus solution. For the experiments shown in Figure 1C, 2 μg of pNLCSFV3 or its derivatives (pNLCSFV3-4A, pNLCSFV3-5A, or pNLCSFV3-4A5A) was cotransfected with 100 ng of flag-tagged APOBEC3D, APOBEC3F, or APOBEC3G expression plasmid into 293T cells. For the experiments shown in Figures 2E and 2F, 500 ng of pNLCSFV3*Δvif* and 500 ng of Vif expression plasmids (see below) were cotransfected with 50 ng of flag-tagged APOBEC3F or APOBEC3G expression plasmid [17] into 293T cells. For the experiments shown in Figure 4G, 2 μg of pNLCSFV3*Δvif* was cotransfected with 100 μg of flag-tagged APOBEC3D, APOBEC3F, or APOBEC3G expression plasmid into 293T cells. The virus solutions were prepared as described above. Then, the virus solutions were treated with DNase I (50 unit; Takara) at 37°C for 1 h and inoculated into TZM-bl cells. The infected TZM-bl cells were harvested at 18 h postinfection and DNA was extracted as described below. Western blotting was performed as previously described [23,25], and anti-Vif antibody (clone #2221; obtained through the NIH AIDS Research and Reference Reagent program) and anti-α-Tubulin (TUBA) monoclonal antibody (DM1A; Sigma) were used. To quantify the infectivity of virus solution, TZM-bl assay was performed as previously described [23,25]. MaRBLE assay was performed as previously described [68] with minor modifications. Briefly, the virus solutions (normalized to the amount of p24 antigen) were inoculated into X4-MaRBLE or R5-MaRBLE cells (1×10^5 cells). At 72 h postinfection, the cells were harvested, and the luciferase activity was measured as previously described [71].

PCR, RT-PCR, and real-time RT-PCR

DNA and RNA were extracted from the splenic human MNCs at 6 wpi or infected TZM-bl cells as previously described [23,25]. cDNA was prepared by using SuperScript III reverse transcriptase (Life technologies) with DNase I (Life technologies), RNaseOUT (Life technologies), and random primers according to the manufacture's procedure. To amplify *vif* ORF (Figures 2E, 2F, and S3), RT-PCR was performed by using PrimeSTAR GXL DNA polymerase (Takara) according to the manufacture's protocol, and the following primers were used: Vif-fwd (4929–4948), 5′-gtt tgg aaa gga cca gca aa-3′; and Vif-rev (5703–5722), 5′-gcc caa gta tcc ccg taa gt-3′. To analyze the sequence of full-length proviral DNA (Figures 4B–4E and S6), PCR was performed by using *Pfu* Ultra II DNA polymerase (Stratagene) according to the manufacture's protocol, and the following primers were used: 5′ region (475–1698, 1,224 bp), 5LTRF#1 (455–474), 5′-ggt ctc tct ggt tag acc ag-3′; and 5LTRR#1 (1699–1718), 5′-gaa gct tgc tcg gct ctt ag-3′; 5′/central region (1342–3530, 2,189 bp), 5F#7 (1322–1341), 5′-gag cca ccc cac aag att ta-3′; and 5/cR#7 (3531–3550), 5′-tgc ccc tgc ttc tgt att tc-3′; central/3′ region (3420–5888, 2,469 bp), C#3F (3400–3419), 5′-ggg gaa cca aag cac taa ca-3′; and 5R#1 (5889–5913), 5′-ttt aca ata gca att ggt aca agc a-3′; 3′ region (5453–9526, 4,074 bp), 3F#2 (5428–5452), 5′-agt cct agg tgt gaa tat caa gca g-3′; and 3LTRR#1 (9527–9547), 5′-ctg gtc taa cca gag aga cc-3′. The products of PCR and RT-PCR were cloned into pCRII-blunt-TOPO by using Zero

blunt TOPO PCR cloning kit (Life technologies) according to the manufacture's protocol. To prepare the expression plasmids of the Vif mutants (Figures 2E and 2F), the pCRII-blunt-TOPO containing *vif* ORFs were digested with *EcoR*I and blunted. The obtained DNA fragments containing *vif* ORF were subcloned into the *Hpa*I site of pDON-AI (Takara). Real-time RT-PCR was performed as previously described [23]. Briefly, *APOBEC3D, APOBEC3F, APOBEC3G* [43] and *IFNB* [72] were amplified by using the primers previously reported. The primers for *GAPDH* were purchased from Life technologies. The expression levels of *APOBEC3D, APOBEC3F*, and *APOBEC3G* (Figure 3C) were standardized as previously described [42,43]. To construct pNLCSFV3 G24R and E25K (Figure 6C), the DNA sequences containing *env* G24R or E25K mutations were digested with *Mlu*I and *Xba*I, and the resultant DNA fragments were subcloned into the *Mlu*I-*Xba*I site of pNLCSFV3.

SGS assay

SGS assay was performed as previously described [50]. Briefly, viral RNA was extracted from the plasma (100 μl) of infected mice at 6 wpi by using QIAamp viral RNA mini kit (Qiagen), and cDNA was prepared as previously described [50].

Sequencing PCR

Sequencing PCR was performed as previously described [23], and the sequence data were analyzed by Seqscape software v2.5 (Applied Biosystems) and Sequencher software (Hitachi). To analyze the sequence of *vif* (Figures 2E, 2F and S3), M13 primers were used. To analyze the sequence of full-length proviral DNA (Figures 4B–4G and S6), M13 primers and the following primers were used: 5#6 (1609–1633), 5′-gta aga atg tat agc cct acc agc a-3′; Cl#1 (2178–2197), 5′-cag gtt tgg gga aga gac aa-3′; C#2 (2700–2719), 5′-ggg cct gaa aat cca tac aa-3′; C#4 (4004–4023), 5′-ttt gca gga ttc ggg att ag-3′; C#5 (4499–4518), 5′-agc aga gac agg gca aga aa-3′; C#6 (5058–5077), 5′-ggt gat gat tgt gtg gca ag-3′; 3#1 (5960–5979), 5′-gca tct cct atg gca gga ag-3′; 3#2 (6651–6660), 5′-gcg gga gaa tga taa tgg ag-3′; 3#3 (7315–7334), 5′-ccc aga aat tgt aac gca ca-3′; 3#4 (7947–7966), 5′-gaa tcc tgg ctg tgg aaa ga-3′; 3#5 (8511–8530), 5′-gct acc acc gct tga gag ac-3′; and 3#6 (8969–8988), 5′-gga gga aga ggt ggg ttt tc-3′. For SGS assay (Figure 5), direct sequencing was performed by using the primers used in the 2nd SGS PCR and the primers 3#2 and 3#3.

Semiquantitative 3D-PCR

Semiquantitative 3D-PCR (Figure 4A) was performed as previously described [48]. Briefly, we used the following primers according to the previous report [48]: 1st-fwd (2723–2746), 5′-tcc art att trc cat aaa raa aaa-3′; 1st-rev (3575–3598), 5′-tty aga ttt tta aat ggy tyt tga-3′; 2nd-fwd (3023–3049), 5′-aat att cca rtr tar cat rac aaa aat-3′; and 2nd-rev (3561–3586), 5′-aat ggy tyt tga taa att tga tat gt-3′. The 1st PCR products were quantified, and constant amounts were used for secondary PCR over 87.5 to 77.0°C range of denaturation temperatures. The 2nd PCR products were run on agarose gels and were visualized by staining with ethidium bromide.

Sequence data analysis and bioinformatics

To analyze the diversity of *vif* in HIV-1 group M (Figure 1A), we obtained 7,118 *vif* ORF sequences registered in Los Alamos HIV sequence database (http://www.hiv.lanl.gov). The 7,118 datasets were aligned by using ClustalW [73] implemented in MEGA 5.1 software [74], and the logo plot shown in Figure 1A was generated by using WebLogo 3 (http://weblogo.threeplusone.com/). To analyze the effect of G-to-A mutation in proviral DNA (Figures 4D and 4E), the sequences of all viral genes (*gag, pol, vif, vpr, tat, rev, vpu, env, and nef*) were obtained from the sequence of full-length proviral DNA (Figure 4B), and the codon-based alignments were constructed using a Gene Cutter tool from the Los Alamos HIV sequence database (http://www.hiv.lanl.gov/content/sequence/GENE_CUTTER/cutter.html). The effect of G-to-A mutation in *env* ORF of viral RNA (Figure 5B) was also analyzed as descried above. To analyze APOBEC3-mediated mutations (Figures 4F and 4G), hypermut 2.0 (http://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html) was used. The *env* ORF sequences obtained by SGS (see above) were aligned by using ClustalW [73] implemented in MEGA 5.1 software [74]. The sequence of WT NLCSFV3 *env* was used as outgroup. The best fitting substitution model was determined using jModelTest 2.1.3. [75]. The Akaike information criterion (AIC) implemented in jmodeltest-2.1.3 selected GTR+I+G as the best-fit. Since this model is not available in MEGA 5.1 software, the TrN+I+G [76], the second best-fit model, was used in further analyses. Genetic distances among *env* ORF sequences (Figure 5D) were calculated with MEGA 5.1 software under the Tamura-Nei model [76]. ML phylogenetic trees (Figure 5C) were reconstructed using PhyML-3.1 under TN93 model [76] with 1,000 bootstrap resamplings [77]. The *env* V3 sequences determined by SGS were used for the genotypic coreceptor usage prediction based on an algorithm, geno2pheno coreceptor [54]. The original g2p coreceptor model was selected, and the sequences below the 10% false-positive rate cutoff were defined as putative CXCR4-tropic viruses (Figure 6B). The major drug resistance sites, which are potentially induced by APOBEC3D and APOBEC3F (Table S1), were determined based on the current IAS-USA lists [78]. The sequence of HIV-1 strain HXB2 (Genbank accession number: FB707281) were used as reference.

3D structure of Vif

The 3D structure of Vif (Figure 1B) was generated on PyMOL v1.6 (http://www.pymol.org/) with the crystal structure of Vif-CBFβ-CUL5-ELOB-ELOC complex (PDB code: 4N9F) [79].

Calculation of AUC of VL

The AUC (Figure 2C) was calculated from the VL data using the trapezoidal rule. For example, let us define that $V(t)$ is a VL at time t. Then the AUC from 0 to 6 wpi is calculated as follows:

$$\int_0^6 V(s)ds = \int_0^1 V(s)ds + \int_1^2 V(s)ds + \dots + \int_5^6 V(s)ds \approx \sum_{w=1}^{6} \{w-(w-1)\}\frac{V(w-1)+V(w)}{2} = \sum_{w=1}^{6}\frac{V(w-1)+V(w)}{2}.$$

Estimation of virus replication rate

To quantify the dynamics during acute virus infection (Figure 2D), we used a recently developed model describing the loss of target cells phenomenologically as follows:

$$T(t) = T(0)\exp(\Delta_1 t), \text{ for } t < t*, \quad (1)$$

$$T(t) = T(0)\exp(\Delta_1 t*)\exp\{\Delta_2(t-t*)\}, \text{ for } t > t*, \quad (2)$$

$$\frac{dI(t)}{dt} = \beta T(t)V(t) - \delta I(t), \quad (3)$$

$$\frac{dV(t)}{dt} = pI(t) - cV(t). \quad (4)$$

Modeling HIV infection it is well accepted to make the quasi-steady assumption, $dV(t)/dt = 0$, and to write that $V(t) = p^*I(t)$, where p^* is a scaled production parameter ($p^* = p/c$). Because we are fitting VLs, $V(t)$, rather than number of infected cells, $I(t)$, we substitute $I(t) = V(t)/p^*$ into Eq. (3) to obtain

$$\frac{dV(t)}{dt} = \{r* T(t) - \delta\}V(t), \quad (5)$$

where $r^* = p^*\beta$ is the viral replication rate per target cell, and δ remains the death rate of infected cells. Because the number of target cells seems to decrease exponentially in phases during the acute phase of several virus infections, we approximated the dynamics of target cells by a piece-wise exponential function. The parameters Δ_1 and Δ_2 represent the two daily loss rates of target cells, and t^* represents the time at which the function switches slope. Because Eqs. (1), (2), and (5) define a non-autonomous linear differential equation, we derived the following analytical solution describing the acute phase of virus infections:

$$V(t) = V(0)\exp\left\{\frac{r* T(0)[\exp(\Delta_1 t) - 1]}{\Delta_1} - \delta(t)\right\}, \text{ for } t < t*, \quad (6)$$

$$V(t) = \quad (7)$$
$$V(t*)\exp\left\{\frac{r* T(t*)[\exp(\Delta_2(t-t*)) - 1]}{\Delta_2} - \delta(t-t*)\right\}, \text{ for } t > t*.$$

We employed the solution of Eqs. (1), (2), (6), and (7) to fit the 6-week time courses of VLs and target cells as shown in Figures 2 and S2 (using the FindMinimum package of *Mathematica 9.0* to minimize the sum of squared residuals). This model has 6 parameters: $T(0)$, Δ_1, Δ_2, $V(0)$, r^*, and δ. The first 3 parameters, $T(0)$, Δ_1, and Δ_2, are estimated from the observed number of peripheral $CD4^+$ T cells per ml of blood. For the latter 3 parameters, $V(0)$, r^*, and δ, we fix $\delta = 1$ per day [80,81], because this is general estimate for the death rate of productively infected cells. The initial value of VL, $V(0)$, was set to the detection limit of the assay (800 copies/ml plasma). The replication rate, r^*, was estimated from the data.

Statistical analysis

Data were presented as averages ± SEMs or SDs. Statistical differences were determined by Student's t test (Figures 1C, 2A–2D, 3C, and 6C), Chi-square test for independence (Figures 4D–4G, 5B, 6B, and S5), and Fisher's exact test (Figure 5C). To determine statistically significant correlation (Figures 3D and S4), Pearson correlation coefficient (r) was applied.

Accession numbers

The GenBank (http://www.ncbi.nlm.nih.gov/genbank/) accession numbers for the genes mentioned in the text are as follows: *APOBEC3D* (NM_152426), *APOBEC3F* (NM_145298), *APOBEC3G* (NM_021822), *IFNB* (NM_002176), and *GAPDH* (NM_002046).

Online supplemental material

Figure S1 shows the anti-viral activity of WT and catalytically inactive APOBEC3 proteins *in vitro*. Figure S2 shows the dynamics of WT HIV-1 and HIV-1 *vif* mutants infection in humanized mice at higher doses. Figure S3 shows the summary of mutations in *vif* ORF. Figure S4 shows the correlation of *APOBEC3* and *IFNB* expressions. Figure S5 shows the statistical analyses on the preferential G-to-A mutation sites. Figure S6 shows the summary of mutations in the proviral DNA of infected humanized mice. Figure S7 shows the raw data of SGS assay. Figure S8 shows the extent of mutation in each amplicon of viral *env*. Table S1 shows putative drug-resistance mutations potentially induced by APOBEC3D and APOBEC3F. Table S2 shows the list of humanized mice used in this study.

Supporting Information

Figure S1 Anti-viral activity of WT and mutated APOBEC3 proteins *in vitro*. Two micrograms of pNLCSFV3*Δvif* was cotransfected with 100 ng of flag-tagged APOBEC3 expression plasmid into 293T cells. (A) Western blotting. The input of cell lysate was standardized to α-Tubulin (TUBA), and representative results are shown. (B) TZM-bl assay. The infectivity of released virus was determined by using TZM-bl cells. The infectivity of each virus is normalized to the value of no APOBEC3. The assay was performed in triplicate. The data represents average with SD.
(TIF)

Figure S2 Dynamics of WT HIV-1 and HIV-1 *vif* mutants infection in humanized mice at higher doses. (A and B) Virus solutions containing 50 ng (A; 4A HIV-1 [n = 8], 5A HIV-1 [n = 8], and 4A5A HIV-1 [n = 6]) or 500 ng (B; 4A HIV-1 [n = 5], 5A HIV-1 [n = 8], and 4A5A HIV-1 [n = 6]) p24 antigens were intraperitoneally inoculated into humanized mice. The amount of viral RNA in plasma (*top*) and the level of peripheral $CD4^+$ T cells ($CD45^+$ $CD3^+$ $CD4^+$ cells) (*bottom*) were analyzed at 0, 1, 2, 3, 5, and 6 wpi. The averages are shown in circles with SEMs, and the values from each mouse are shown by line. In panel A, the detection limit of HIV-1 RNA is 800 copies/ml plasma.
(TIF)

Figure S3 Mutations in *vif* ORF. The *vif* ORFs (5041–5619, 579 bases) of viral RNA in the spleen of infected mice (WT, n = 82 from 3 mice; 4A, n = 320 from 7 mice; 5A, n = 132 from 3 mice; and 4A5A, n = 59 from 1 mouse) were sequenced. (A) The mutation matrix (*left*) and the pie chart of G-to-A mutation (*right*) are shown. In the *right* panel, the diameters of pie charts represent the percentage of G-to-A mutations in total mutations. (B) The extent of mutation in each amplicon of *vif* ORF sequences. The numbers of total mutations (*top*, gray), G-to-A mutations (*upper middle*, black), GA-to-AA mutations (*lower middle*, red), and GG-to-AG mutations (*bottom*, blue) within each amplicon are respectively shown. (C) The *vif* amplicons harboring nonsynonymous mutations are summarized. "X" means stop codon mutation. The *vif* ORFs indicated by asterisks were used for the functional assay, and the results are shown in Figures 2E and 2F. In the panel of 4A5A, the amplicon harboring both R132K and I124M mutations (indicated by double

daggers) were frequently detected and were used for the functional assay (Figures 2E and 2F).
(TIF)

Figure S4 Correlation of *APOBEC3* and *IFNB* expressions. The mRNA expression levels of *APOBEC3D*, *APOBEEC3F*, *APOBEC3G*, and *IFNB* in the splenic human CD4$^+$ T cells of humanize mice (n = 73) were measured by real-time RT-PCR. The expression level of each gene was normalized to that of *GAPDH* and was shown as relative expression. The correlation between each *APOBEC3* (A) and between *APOBEC3* (x-axes) and *IFNB* (y-axes) (B) are respectively shown. The lines represent exponential approximation. Pearson correlation coefficient (r) was adopted to determine statistically significant correlation between each value.
(TIF)

Figure S5 Statistical analyses on the preferential G-to-A mutation sites. The detected G-to-A mutation sites in the proviral DNA of *vif*-mutated HIV-1-infected mice (A) and *in vitro* experiments (B) were classified according to the nucleotides positioned between −5 to +5 from the detected G-to-A mutation sites (position 0; see also Figures 4F and 4G). Statistical differences at each position were determined by Chi-square test for independence, and the P values at each position is shown in y-axes. Statistically significant differences ($P<0.001$) are shown with asterisks.
(TIF)

Figure S6 Mutations in the proviral DNA of infected humanized mice. The percentages of mutations in each site are summarized.
(TIF)

Figure S7 Raw data of SGS assay. The *env* ORF (6221–8782, 2,562 bases) of viral RNA in the plasma of infected mice (WT, n = 73 from 2 mice; 4A, n = 91 from 3 mice; 5A, n = 68 from 2 mice; and 4A5A, n = 33 from 1 mouse) were analyzed by SGS assay, and the raw data are shown.
(TIF)

Figure S8 The extent of mutation in each amplicon of viral *env* in plasma. The numbers of total mutations (*top*, gray), G-to-A mutations (*upper middle*, black), GA-to-AA mutations (*lower middle*, red), and GG-to-AG mutations (*bottom*, blue) within each amplicon are respectively shown.
(TIF)

Table S1 Putative drug-resistance mutations potentially induced by APOBEC3D and APOBEC3F. The table provides the drug-resistance mutation sites potentially induced by APOBEC3D and APOBEC3F.
(PDF)

Table S2 Humanized mice used in this study. A full list of the 82 humanized mice used in this study.
(PDF)

Acknowledgments

We would like to thank Dr. John Coffin (Tufts University) for lively discussion and providing critical comments, Dr. Fengrong Ren (Tokyo Medical and Dental University) for supporting genetic distance analyses, Dr. Tobias Paprotka (NCI-Frederick, NIH) for helpful suggestion, Dr. Wataru Sugiura (National Hospital Organization Nagoya Medical Center) for providing experimental materials, Drs. Yoshinori Fukazawa (Oregon Health and Science University), Keiko Okano (Kyoto University Research Administration Office) and Seiga Ohmine (Mayo Clinic) for proofreading the manuscript, the colleagues belonging to iGEM Kyoto (Sakaya Aoki, Yoshitaka Hirano, Akiho Kojima, Yusuke Komoto, Mina Noura, Kenji Okumura, Toshiro Shimizu, Masako Terasaka, Moe Yanagi, and Mitsuaki Yoshida) for their experimental support, and Ms. Kotubu Misawa for the dedicated support.

Author Contributions

Conceived and designed the experiments: KS TI VKP YKo. Performed the experiments: KS JST NM TI TK YKi. Analyzed the data: KS JST NM TI TK YKi SI. Contributed reagents/materials/analysis tools: NM SI ATK WSH KA MI DSA. Wrote the paper: KS VKP YKo.

References

1. Conticello SG (2008) The AID/APOBEC family of nucleic acid mutators. Genome Biol 9: 229.
2. Conticello SG, Langlois MA, Yang Z, Neuberger MS (2007) DNA deamination in immunity: AID in the context of its APOBEC relatives. Adv Immunol 94: 37–73.
3. Teng B, Burant CF, Davidson NO (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science 260: 1816–1819.
4. Harris RS, Liddament MT (2004) Retroviral restriction by APOBEC proteins. Nat Rev Immunol 4: 868–877.
5. Zhang J (2003) Evolution by gene duplication: an update. Trends Ecol Evol 18: 292–298.
6. Sawyer SL, Emerman M, Malik HS (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol 2: E275.
7. Stenglein MD, Burns MB, Li M, Lengyel J, Harris RS (2010) APOBEC3 proteins mediate the clearance of foreign DNA from human cells. Nat Struct Mol Biol 17: 222–229.
8. Burns MB, Temiz NA, Harris RS (2013) Evidence for APOBEC3B mutagenesis in multiple human cancers. Nat Genet 45: 977–983.
9. Shinohara M, Io K, Shindo K, Matsui M, Sakamoto T, et al. (2012) APOBEC3B can impair genomic stability by inducing base substitutions in genomic DNA in human cells. Sci Rep 2: 806.
10. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, et al. (2013) APOBEC3B is an enzymatic source of mutation in breast cancer. Nature 494: 366–370.
11. Sheehy AM, Gaddis NC, Choi JD, Malim MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 418: 646–650.
12. Izumi T, Shirakawa K, Takaori-Kondo A (2008) Cytidine deaminases as a weapon against retroviruses and a new target for antiviral therapy. Mini Rev Med Chem 8: 231–238.
13. Dang Y, Wang X, Esselman WJ, Zheng YH (2006) Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. J Virol 80: 10522–10533.
14. Liddament MT, Brown WL, Schumacher AJ, Harris RS (2004) APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 *in vivo*. Curr Biol 14: 1385–1391.
15. Refsland EW, Hultquist JF, Harris RS (2012) Endogenous origins of HIV-1 G-to-A hypermutation and restriction in the nonpermissive T cell line CEM2n. PLoS Pathog 8: e1002800.
16. Miyagi E, Brown CR, Opi S, Khan M, Goila-Gaur R, et al. (2010) Stably expressed APOBEC3F has negligible antiviral activity. J Virol 84: 11067–11075.
17. Chaipan C, Smith JL, Hu WS, Pathak VK (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.
18. Pillai SK, Wong JK, Barbour JD (2008) Turning up the volume on mutational pressure: is more of a good thing always better? (A case study of HIV-1 Vif and APOBEC3). Retrovirology 5: 26.
19. Casartelli N, Guivel-Benhassine F, Bouziat R, Brandler S, Schwartz O, et al. (2010) The antiviral factor APOBEC3G improves CTL recognition of cultured HIV-infected T cells. J Exp Med 207: 39–49.
20. Jern P, Russell RA, Pathak VK, Coffin JM (2009) Likely role of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance. PLoS Pathog 5: e1000367.
21. Wood N, Bhattacharya T, Keele BF, Giorgi E, Liu M, et al. (2009) HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. PLoS Pathog 5: e1000414.
22. Nie C, Sato K, Misawa N, Kitayama H, Fujino H, et al. (2009) Selective infection of CD4$^+$ effector memory T lymphocytes leads to preferential depletion of memory T lymphocytes in R5 HIV-1-infected humanized NOD/SCID/IL-2Rγ^{null} mice. Virology 394: 64–72.