

RESEARCH ARTICLE

Role of *APOBEC3F* Gene Variation in HIV-1 Disease Progression and Pneumocystis Pneumonia

Ping An^{1*}, Sudhir Penugonda², Christian W. Thorball^{3,4}, Istvan Bartha^{3,4}, James J. Goedert⁵, Sharyne Donfield⁶, Susan Buchbinder⁷, Elizabeth Binns-Roemer¹, Gregory D. Kirk⁸, Wenyan Zhang⁹, Jacques Fellay^{3,4}, Xiao-Fang Yu^{9,10}, Cheryl A. Winkler^{1*}

1 Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, United States of America, **2** Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States of America, **3** School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, **4** Swiss Institute of Bioinformatics, Lausanne, Switzerland, **5** Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, United States of America, **6** Rho, Inc., Chapel Hill, North Carolina, United States of America, **7** San Francisco Department of Public Health, San Francisco, California, United States of America, **8** Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, Maryland, United States of America, **9** Institute of Virology and AIDS Research, First Hospital of Jilin University, Changchun, China, **10** Department of Molecular Microbiology and Immunology, Johns Hopkins School of Public Health, Baltimore, Maryland, United States of America

* ping.an@nih.gov (PA); winklerc@mail.nih.gov (CAW)



OPEN ACCESS

Citation: An P, Penugonda S, Thorball CW, Bartha I, Goedert JJ, Donfield S, et al. (2016) Role of *APOBEC3F* Gene Variation in HIV-1 Disease Progression and Pneumocystis Pneumonia. *PLoS Genet* 12(3): e1005921. doi:10.1371/journal.pgen.1005921

Editor: Giorgio Sirugo, Ospedale San Pietro Fatebenefratelli, ITALY

Received: August 11, 2015

Accepted: February 16, 2016

Published: March 4, 2016

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The Hemophilia Growth and Development Study is funded by the National Institutes of Health, National Institute of Child Health and Human Development (R01-HD-41224, JJG). AIDS Linked to Intravenous Experience is supported by National Institutes of Health (grants R01-DA-04334 and R01-DA-12568, GDK). The MACS is funded primarily by the National Institute of Allergy and Infectious Diseases (NIAID), with additional co-funding from the National Cancer Institute (NCI), the National Institute

Abstract

Human APOBEC3 cytidine deaminases are intrinsic resistance factors to HIV-1. However, HIV-1 encodes a viral infectivity factor (Vif) that degrades APOBEC3 proteins. *In vitro* APOBEC3F (A3F) anti-HIV-1 activity is weaker than A3G but is partially resistant to Vif degradation unlike A3G. It is unknown whether A3F protein affects HIV-1 disease *in vivo*. To assess the effect of A3F gene on host susceptibility to HIV-1 acquisition and disease progression, we performed a genetic association study in six well-characterized HIV-1 natural cohorts. A common six-Single Nucleotide Polymorphism (SNP) haplotype of A3F tagged by a codon-changing variant (p. I231V, with allele (V) frequency of 48% in European Americans) was associated with significantly lower set-point viral load and slower rate of progression to AIDS (Relative Hazards (RH) = 0.71, 95% CI: 0.56, 0.91) and delayed development of pneumocystis pneumonia (PCP) (RH = 0.53, 95% CI: 0.37–0.76). A validation study in the International Collaboration for the Genomics of HIV (ICGH) showed a consistent association with lower set-point viral load. An *in vitro* assay revealed that the A3F I231V variant may influence Vif mediated A3F degradation. Our results provide genetic epidemiological evidence that A3F modulates HIV-1/AIDS disease progression.

Author Summary

Cytidine deaminases of the human APOBEC3 (A3) gene family serve as intrinsic resistance factors to HIV-1 and other retroviruses. HIV-1 encodes the viral infectivity factor

on Drug Abuse (NIDA), and the National Institute of Mental Health (NIMH). The MACS data in this manuscript were collected by the Multicenter AIDS Cohort Study Collaborative Study Group with centers (Principal Investigators) located at: Johns Hopkins University Bloomberg School of Public Health (Joseph Margolick), U01-AI35042; Northwestern University (Steven Wolinsky), U01-AI35039; University of California, Los Angeles (Roger Detels), U01-AI35040; University of Pittsburgh (Charles Rinaldo), U01-AI35041; the Center for Analysis and Management of MACS, Johns Hopkins University Bloomberg School of Public Health (Lisa Jacobson), UM1-AI35043. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract HHSN26120080001E. This Research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: PA, EBR and CAW are employed by Leidos Biomedical Research, Inc and SD by Rho, Inc. All other authors have declared that no competing interests exist.

(Vif) protein that degrades APOBEC3 proteins via the ubiquitination-proteosomal pathway. APOBEC3F (A3F), unlike APOBEC3G (A3G), is partially resistant to Vif-mediated degradation. The antiviral activity of the A3 family has largely been demonstrated in *in vitro* experiments, and there is mounting evidence that A3G genetic variants influence HIV disease progression. It is not resolved if A3F protein affects HIV disease *in vivo*. To assess the *in vivo* effect of A3F, we performed a genetic association study of genetic variants in *A3F* for their influence on HIV- acquisition and HIV disease progression. A common *A3F* haplotype was associated with a 30% reduced rate of AIDS disease progression, lower set-point viral load and delayed development of pneumocystis pneumonia (PCP) in European Americans. This study provides the first epidemiological evidence that A3F might modify HIV-1/AIDS pathogenesis.

Introduction

The apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3, A3) proteins are a family of cellular cytidine deaminases that defend against a diverse set of retroviruses, endogenous retroelements and DNA viruses, including human immunodeficiency virus type I (HIV-1) [1–5]. Humans A3 proteins are encoded by seven *A3* genes (*A3A*, *A3B*, *A3C*, *A3D*, *A3F*, *A3G*, and *A3H*) tandemly arrayed on chromosome 22. In *in vitro* experimental systems, several members of the human APOBEC3 family are capable of inhibiting HIV-1 replication to some degree (A3G, A3F, A3D and some A3H haplotypes), with A3G and A3F showing evidence of strong inhibitory activity. A3G protein catalyzes deamination of cytosine bases on the DNA minus strand during reverse transcription, inducing guanosine (G)-to-adenosine (A) hypermutation in the HIV-1 provirus [2,6–8]. A3G causes GG-to-AG transitions, while APOBEC3F causes GA-to-AA nucleotide changes. However, HIV-1 encodes an accessory protein, viral infectivity factor (Vif), to counteract APOBEC3 proteins by mediating the proteasomal degradation of A3 proteins [9–13]. It is unknown whether the anti-HIV-1 activity of A3 proteins is completely neutralized by Vif or if A3 still exerts meaningful antiviral effect *in vivo*. While G to A hypermutation of the HIV genome is deleterious to the virus [10], [16–18], a recent *in vitro* study showed sub-lethal levels of A3G induced G to A mutations may contribute to viral diversity, which consequently may contribute to immune escape and drug resistance [14]. How A3 proteins' opposing antiviral and viral mutation properties contribute to HIV-1 pathogenesis *in vivo* is a critical question.

The antiviral strength of A3F is unresolved by *in vitro* experiments as some reported that A3F activity is as strong as A3G [7,15–17] while others indicated that it is weaker [18–21], possibly due to varied experimental settings. A3F is highly expressed in CD4+ T-cells, the cells infected by HIV-1 [22,23]. Unlike A3G, A3F is partially resistant to Vif-mediated degradation [15]. A3F also has a favored deamination sequence target that differs from A3G (5'-TC for A3F and 5'-CC for A3G) [24]. Patterns of HIV DNA hypermutation observed in HIV-1 infected individuals are consistent with the notion that both A3G and A3F might be acting on HIV-1 *in vivo* [25].

Testing the *in vivo* role of intrinsic host restriction factors such as A3 proteins are challenging due to a lack of a good animal model [26,27]. The genetic associations between natural polymorphisms in *A3* genes and resistance to infection to HIV-1 or restriction to HIV-1 disease progression would provide critical evidence supporting an *in vivo* role for A3 proteins in restricting HIV. Genetic variation in *A3G* and *A3H* have been shown to associate with HIV-1 disease progression [28–30], supporting the *in vivo* activity of A3G and A3H's. However, it

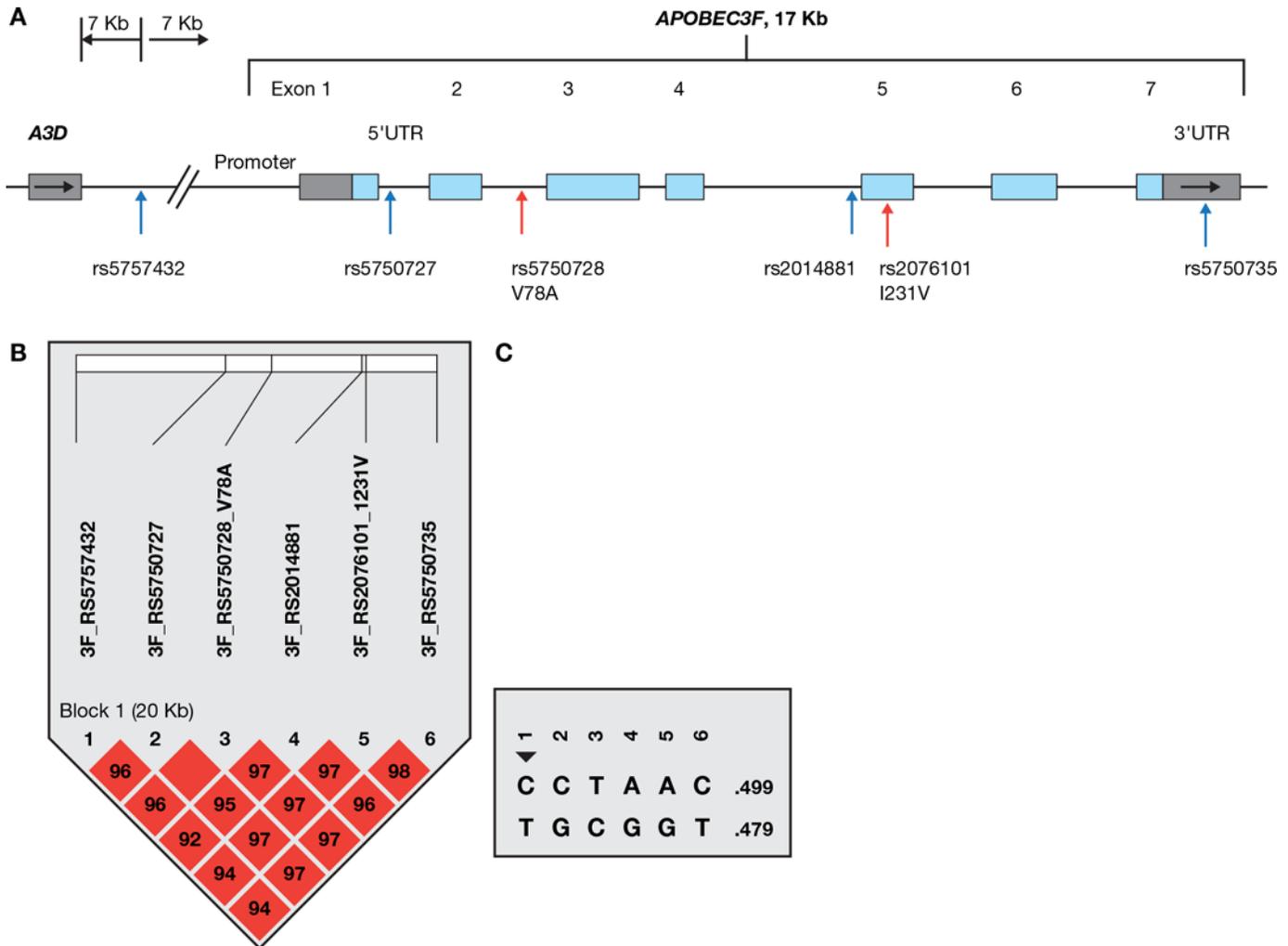


Fig 1. SNPs analyzed in the *A3F* gene. (A) Gene structure and SNP locations. The colored blocks indicate exons, empty blocks, untranslated regions (UTR), horizontal arrows, the direction of transcription and vertical arrows, the positions of SNPs. (B) Linkage disequilibrium matrix of SNPs in the *A3F* gene region in European Americans, as illustrated by Haploview [74]. Red block indicates $D' = 1.0$, and the number in the blocks indicates the value of D' . The linkage disequilibrium block depicted by black triangle was based on the 95% Confidence interval criteria. (C) Haplotypes in the *A3F* gene region in European Americans.

doi:10.1371/journal.pgen.1005921.g001

remains unknown whether *A3F* confers any appreciable effect on HIV acquisition or HIV pathogenesis. We investigated the association of variants in the *A3F* gene with susceptibility to HIV-1 acquisition, HIV-1 viral load, and progression to AIDS in six clinically well-characterized, natural history HIV cohorts.

Results

Linkage disequilibrium, haplotype structure and characteristics of *A3F* variants

A3F is about 17 kb in length, containing 7 exons. Six potentially functional SNPs, including the missense SNP rs2076101 ATC>GTC, coding p.I231V, were genotyped in the AIDS cohorts (Fig 1 and Table 1). The allele rs2076101G (231V) had a frequency of 48% in European

Americans (EA) and 76% in African Americans (AA) in our cohorts. Genotype distributions of the 6 SNPs conformed to Hardy-Weinberg equilibrium expectations ($P > 0.05$) in European Americans and African Americans. The 6 SNPs were in near-absolute linkage disequilibrium (LD) and highly correlated ($D' = 1$ and $r^2 > 0.95$, respectively) in EA. SNP rs2076101A/G (p.I231V) tags the most frequent A3F haplotype comprising the variant alleles of all 6 SNPs (Fig 1); we therefore used rs2076101A/G (p.I231V) to represent the A3F haplotype in the association analyses.

Based on HapMap LD map covering the APOBEC3 gene family region (S1 Fig), A3F, 3G, 3H each gene forms a distinct haplotype block but are not in strong LD with each other, as reported previously [29].

The possible functional consequence of the six SNPs, as evaluated by the Variant Effect Predictor (VEP) function from Ensembl (www.ensembl.org/), is listed in Table 1. The missense SNP rs2076101 (p.I231V) results in a change from isoleucine to valine, has a Polyphen score of 0.90, reflecting a non-conservative change with possible deleterious consequence. SNP rs5750728 is located in the intron of A3F encoding the major transcript; however, the shorter transcript encoding 101 amino acids (NM_001006666) would contain p.V78A, a benign change. Whether this isoform is functional *in vivo* is still unknown. SNP rs5750727 in intron 1 is located within an experimentally identified regulatory feature in CD4+ T-cells (ENSR00001532770, based on Ensembl database).

Impact of A3F 231V haplotype on HIV-1 progression

To assess the impact of the A3F 231V haplotype on disease progression from the estimated date of seroconversion to clinical AIDS, we performed time-to-event analysis for 707 European American seroconverters. The seroconversion date was estimated as the midpoint between the last seronegative and the first seropositive HIV-1 antibody test date (mean seroconversion interval 0.89 years, range 0.06–3.0 years). In the Kaplan-Meier survival curve analysis, p. 231V was associated with delayed progression to clinical AIDS (Fig 2A) ($P = 0.02$, in both log-rank and Wilcoxon test) in the dominant genetic model. To account for covariates (age, sex, and cohort) and genetic confounders (*CCR5* and *HLA* variants), we employed adjusted Cox proportional hazards regression models to evaluate progression rates to clinical AIDS in 707 European American seroconverters. After stratifying by age group, sex and cohort, carriers of one or two copies of p. 231V showed significantly delayed progression to clinical AIDS (dominant genetic model, RH = 0.68, 95% CI, 0.54–0.87, $P = 0.002$); In the model adjusted for the known genetic factors of *HLA* alleles (*HLA-C*, *B27*, *B57*, *B35Px*, *HLA* class I homozygosity, and *CCR5* (p1 and Δ32)[31,32], the association remained significant (RH = 0.71, 95% CI, 0.56–0.91,

Table 1. Characteristics of A3F genetic variants.

SNP rs#	Position on Chr. 22	Location	codon change	Allele change	Predicted SNP function ^a	Polyphen ^b
rs5757432	37758865	Intergenic		C/T	miRNA	
rs5750727	37767356	Intron		C/G	TFBS	
rs5750728	37770095	Intron or alternative exon	V78A	C/T	TFBS Splicing(ESE)	benign
rs2014881	37775326	Intron		A/G		
rs2076101	37775500	exon	I231V	A/G	Splicing(ESE), nsSNP	possibly damaging
rs5750735	37779601	3'UTR		C/T	miRNA	

^aPredicted by SNPinfo web server (<http://snpinfo.niehs.nih.gov/>); TFBS, transcription factor-binding site; miRNA, MicroRNA-binding sites; nsSNP, non-synonymous SNP; ESE, exonic splicing enhancer.

^bPredicted by Variant Effect Predictor from Ensembl (www.ensembl.org/)

doi:10.1371/journal.pgen.1005921.t001

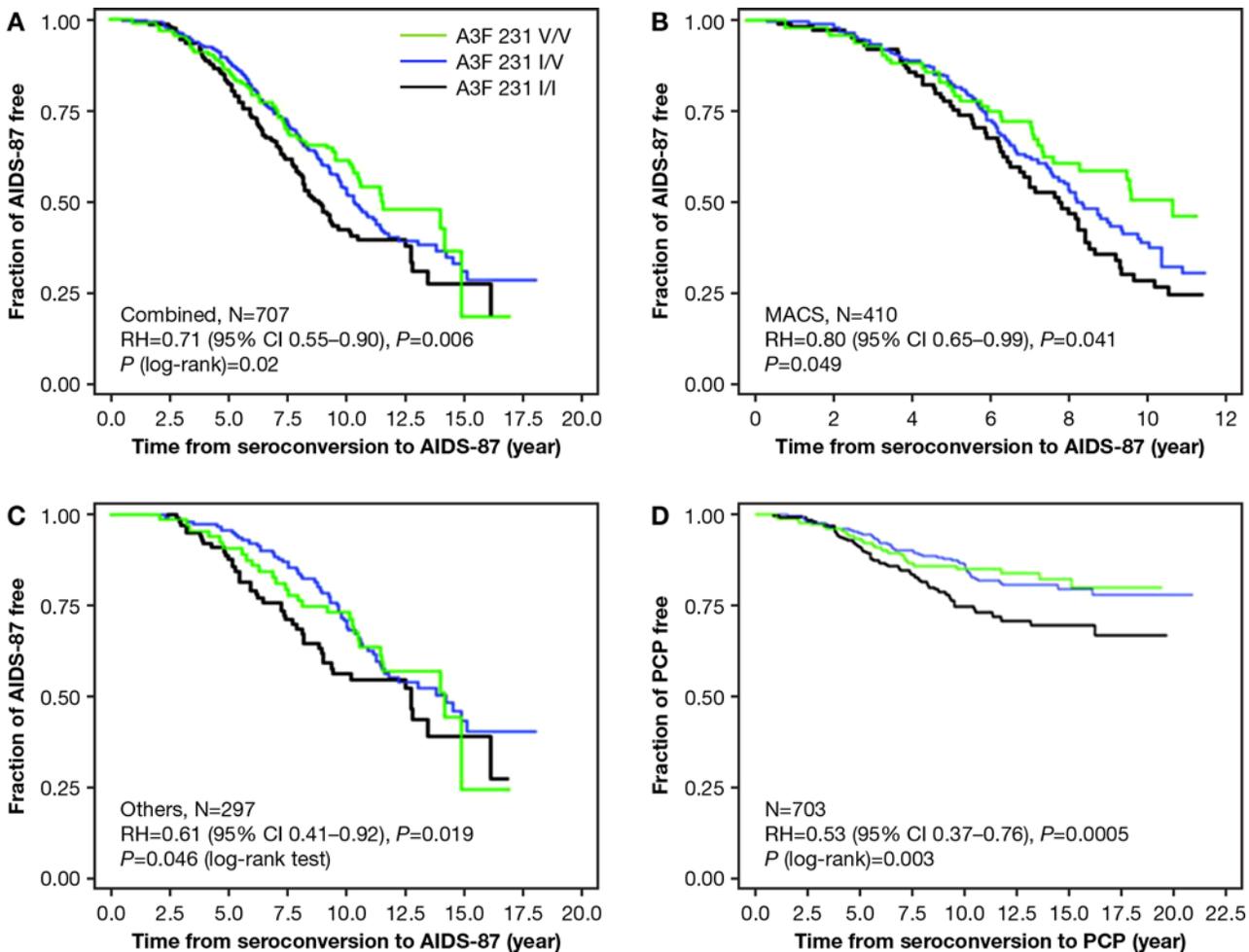


Fig 2. Genetic effect of A3F 231V haplotype on AIDS progression. Kaplan-Meier survival curves of A3F-231V genotype carriers progression to AIDS since HIV-1 infection in the (A) combined cohorts. (B) MACS cohort. (C) other cohorts. (D). Kaplan-Meier survival curves of A3F-231V genotype carriers for progression to PCP since HIV-1 infection in 703 seroconverter European Americans. RH and adjusted P values were obtained from a Cox proportional hazards model. P values for survival curves were obtained from a log-rank test.

doi:10.1371/journal.pgen.1005921.g002

$P_{adj.} = 0.007$, Table 2). Further adjustment for potential population stratification using the first 2 eigenvalues led to similar result (RH = 0.71, 95% CI, 0.55–0.90; $P_{eigen} = 0.006$). The A3F association was not attenuated (RH = 0.68, 95% CI, 0.51–0.90) with additional adjustment for the

Table 2. Impact of A3F 231V haplotype on AIDS progression assessed by Cox proportional hazards regression model.

Population	Disease outcome	no. total/events	Adjustment	RH (95% CI)	P
European American	AIDS 1987	707/318	unadjusted	0.68 (0.54–0.87)	0.002
			adjusted ^a	0.71 (0.55–0.90)	0.006
	PCP	703/134	unadjusted	0.53 (0.37–0.76)	0.0005
Africa American	AIDS 1987	281/48	unadjusted	0.39 (0.12–1.31)	0.12
			adjusted ^b	0.34 (0.10–1.18)	0.09

Dominant genetic model (231V/V or V/I vs. I/I) was tested and was stratified by sex, age and cohort.

^a adjusted for covariates HLA homozygosity, HLA-C, HLA-B*57, HLA-B*35PX, HLA-B*27, CCR5-Δ32, CCR5-59029 and the first 2 eigenvalues

^b adjusted for covariates HLA homozygosity, HLA-B*57, HLA-B*35PX, HLA-B*27, and CCR5-59029.

doi:10.1371/journal.pgen.1005921.t002

known HIV progression modifiers *A3B* gene deletion [33] and *A3G* 199376C [30]. The effect was not affected by the potential seroconversion estimate errors; RHs were changed from 0.71 to 0.72, 0.72, respectively, when seroconversion intervals were shortened to < 2.0 (n = 649) or to < 1.5 years (n = 585). The additive genetic model also showed significant association (RH = 0.80, 95% CI, 0.68–0.95, $P_{eigen} = 0.009$). Bayesian analysis using the Gibbs sampler with 50,000 iterations provided the same results, indicating the p values are robust. Separated analyses conducted in the MACS cohort alone (Fig 2B) or in the other cohorts (Fig 2C) showed significant association in the same direction. These results provide evidence that the haplotype carrying the 6 variant alleles including p. 231V was associated with a 30% reduced rate of AIDS disease in European Americans. In 281 African Americans seroconverters, p. 231V trended in the same direction in a Cox model analysis (RH = 0.35, 95% CI, 0.10–1.19, $P_{adj.} = 0.09$; RH = 0.30, 95% CI, 0.07–1.37, $P = 0.12$ with further adjustment of population stratification).

The protective effect of p. 231V for AIDS progression was also observed in a defined disease category analysis (S2 Fig), which allows the addition of seroprevalent patients who had not progressed to AIDS for >10 years (time from seroconversion to AIDS) to the slow progressor category of seroconverters; only seroconverters were included in the fast category (≤ 10 years) to avoid frailty bias. The prevalence of AIDS protective genotypes 231V/V and 231 I/V was elevated in the slow progressors (Odds ratio (OR) = 0.62, $P = 0.004$, χ^2 test, S2 Fig). The results of both the survival (Fig 2A and Table 2) and defined disease category analyses (S2 Fig) support a strong dominant 231V association with delayed progression to clinical AIDS outcome.

A3F 231V haplotype association with pneumocystis pneumonia (PCP)

Multiple different AIDS-defining conditions including opportunistic infections and malignancies are AIDS-defining [34]. In an explanatory data analysis, we tested whether the common AIDS defining condition, pneumocystis pneumonia was influenced by p.231V. In a Kaplan-Meier survival curve analysis of 703 seroconverters, carriers of 1 or 2 copies of the p. 231V haplotype showed a significantly slower progression to pneumocystis pneumonia, compared with 231I/I carriers (log-rank test, $P = 0.003$; Wilcoxon test, $P = 0.005$, Fig 2D). A Cox model analysis also showed an association of p.231V with delayed progression to PCP (dominant model, RH = 0.53, 95% CI 0.37–0.76, Wald test $P_{dom} = 0.0005$, Table 2; additive model, RH = 0.69, 95% CI 0.54–0.89, $P = 0.004$; both adjusted for age, sex and cohort). A sensitivity test in the MACS cohort revealed no impact of PCP prophylaxis usage on the association. In a sensitivity test performed in 410 MACS seroconverters, adjustment of PCP prophylaxis (n = 138) did not attenuate the association of p.231V with PCP (RH changed from 0.637 to 0.633). These results demonstrated that p.231V has a protective association with PCP development. In 278 African American seroconverters, a nonsignificant protective trend to PCP was seen for p.231V homozygote carriers (OR = 0.55, 95% CI, 0.28–1.12, $P = 0.10$; S3 Fig)

A3F 231V haplotype association with HIV-1 viral load

Among 183 MACS seroconverters with available viral load set-point data, individuals carrying the p.231V (n = 141) had significantly lower mean viral load than those homozygous for the p.231I (n = 42) (mean 4.46 ± 0.69 versus 4.72 ± 0.46 log₁₀ copies/ml; $P = 0.02$). The p.231V carriers were associated with a 0.27 ± 0.65 log₁₀ lower set-point viral load (equivalent of 1.86-fold copies per ml of plasma). For comparison, heterozygosity for *CCR5* $\Delta 32$ [35], had a 0.23 ± 0.65 lower viral load ($P = 0.04$) in this dataset, supporting the validity of the analysis.

To replicate the viral load association result, we obtained the association results of the A3F SNPs with HIV-1 viral load from the International Collaboration for the Genomics of HIV (ICGH) that combined genome-wide SNP data from 25 cohorts worldwide [36]. We analyzed

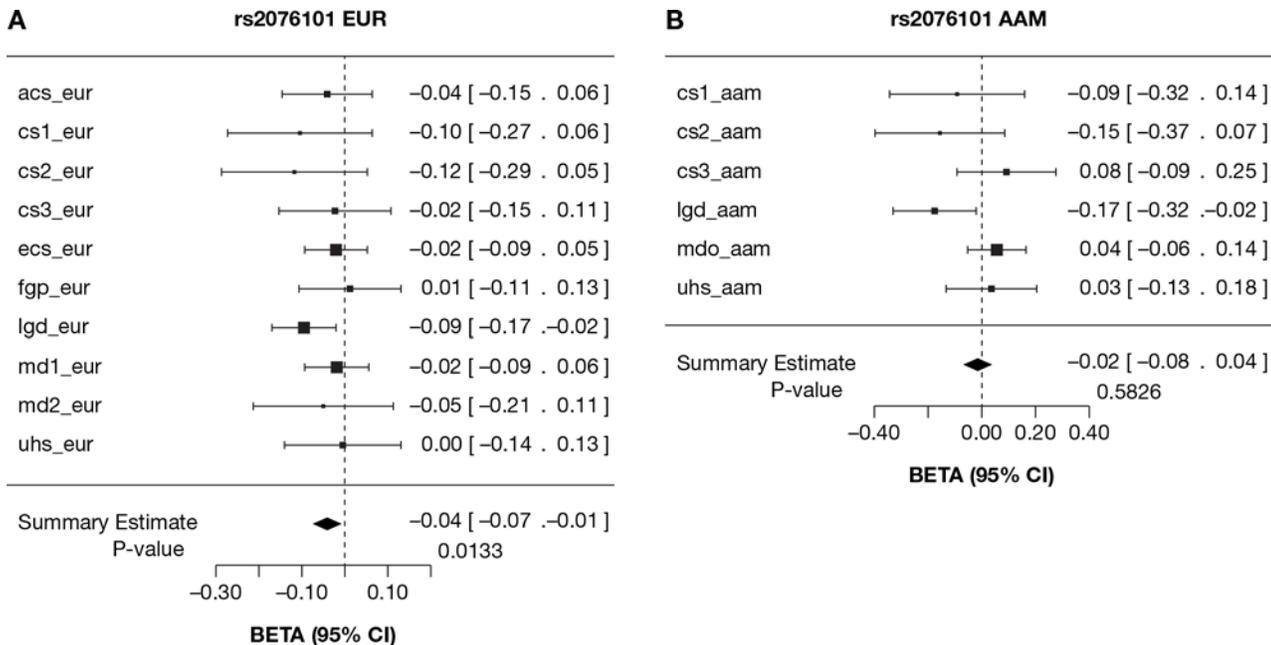


Fig 3. HIV-1 viral load levels of A3F 231V allele carriers in multiple HIV-1 cohorts. Forest plot of effect estimates of HIV-1 viral load for the A3F rs2076101 G allele with 95% confidence intervals per study group (box and whiskers) and after meta-analysis (diamond) in the International Collaboration for the Genomics of HIV (ICGH) consortium [36]. (A) Meta-analyses for European or European-descents (eur). (B) Meta-analyses for African Americans (aam). A description of the study groups and cohorts used in the meta-analysis is presented in S3 Table.

doi:10.1371/journal.pgen.1005921.g003

13 Europe- or USA-based cohorts in ICGH including 6 cohorts used in this study with viral load data available. For independent confirmation, we removed 183 MACS seroconverter subjects used in the above viral load analysis from the ICGH. A meta-analysis of 13 cohorts comprising 10,395 seropositives showed a consistent association of A3F rs2076101 (231V) with lower set-point viral load ($\beta = -0.04$, 95% CI, -0.07, -0.01, $P = 0.01$) in EA (Fig 3A). The association was mainly driven by the cohorts used in this study (comprising cohorts MACS, MHCS, DCG, HGDS, SFCCC and ALIVE), and most other European-ancestry cohorts showed nonsignificant trends in the same direction. In African American subjects, a meta-analysis showed no association with viral load but a significant protective association was seen for the ALIVE cohort (95% CI of β , -0.32, -0.02, Fig 3B). Together, these results indicate a modest but consistent association between p.231V and lower viral load.

Impact of A3F 231V on HIV-1 acquisition

No differences in the A3F p.231V frequencies were observed between African- or European American HIV-1 seronegative groups and HIV-1-infected group (S1 Table), indicating that the A3F genetic variation assessed herein has no obvious effect on susceptibility to HIV-1 acquisition.

A3F rs5750728 alters transcriptional factor binding

The intronic SNP rs5750728 is in near absolute LD with 231V and is predicted to have regulatory function (Table 1). We used electrophoretic mobility shift assays (EMSA) to determine if the variant alters transcription factor binding. The rs5750728 T containing oligonucleotide probe formed a major DNA-protein complex with HeLa cell nuclear extract, which was nearly absent when using rs5750728 C containing oligonucleotide (Fig 4A, lanes 3 and 4). The DNA-protein complex was abolished by the addition of excessive unlabeled competitor probes

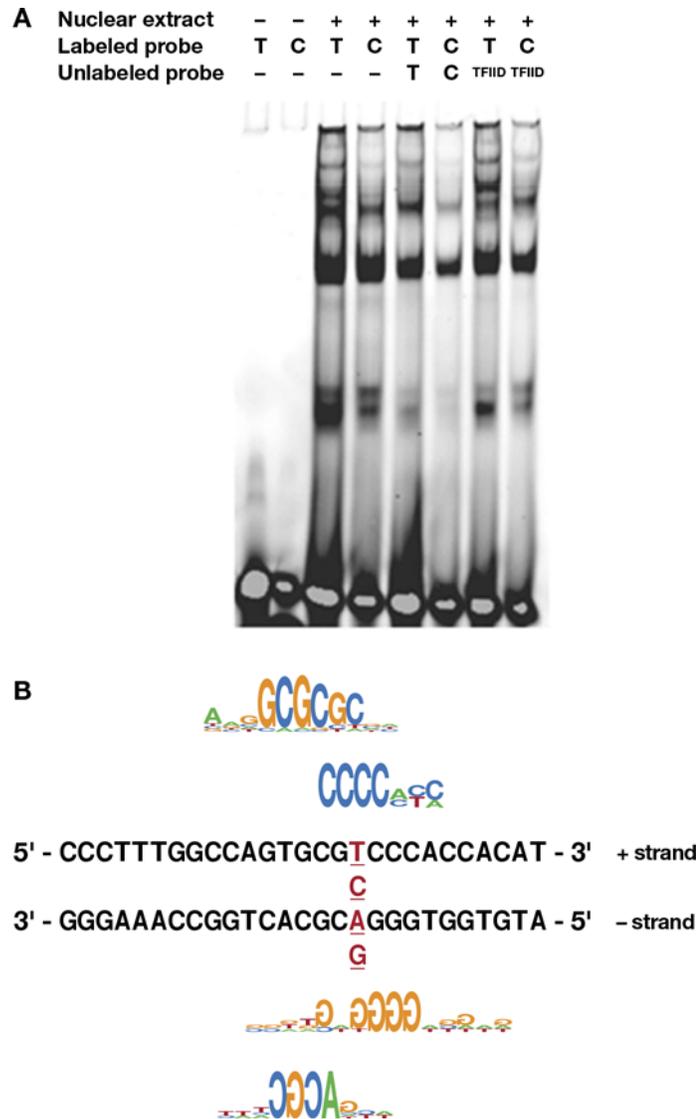


Fig 4. Allelic-specific protein binding of A3F rs5750728T/C. (A) For the EMSA, two A3F DNA fragments with rs5750728 T or C allele were incubated with HeLa nuclear extract. The specificity of the intense bands present in the rs5750728 T lane (lane 3) was demonstrated by adding 200-fold excess of unlabeled T oligo (lane 5) and unrelated TFIID oligos (lane 7) as competitors. (B) Transcription factor binding sites affected by the rs5750728 T/C change based on SNPInspector. The C allele of rs5750728 generates E2FF, ZKSCAN3 (matching the positive (+) strand of the rs5750728 residing A3F DNA fragment), KLFS and loss of WHNF (matching the negative—strand) sites. The rs5750728 site is in red and underlined.

doi:10.1371/journal.pgen.1005921.g004

(Fig 4A, lanes 5 and 6), but was not affected by non-specific competitor probes (Fig 4A, lanes 7 and 8), demonstrating the binding specificity. These results indicate that the rs5750728 C variant allele caused a loss of binding to certain transcription factors, although the identity of the transcription factors is not yet known.

Genomatix software SNPInspector predicted that rs5750728 T/C change generates transcription factor binding sites for E2FF (myc activator/cell cycle regulator), KLFS (Krueppel like transcription factors), and ZKSCAN3 (zinc finger with KRAB and SCAN domains 3), and loss of WHNF (winged helix binding sites) (Fig 4B). Another SNP rs5750727 C/G causes lost site for MRF2 (modulator recognition factor 2).

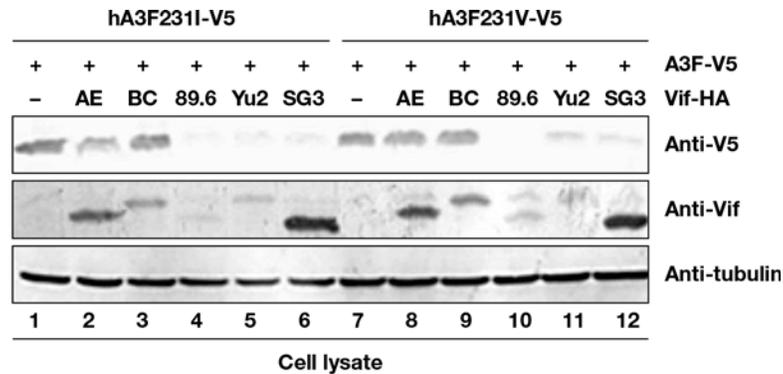


Fig 5. Influence of HIV-1 Vif on A3F variant protein expression. HEK293T cells in 12-well plate were co-transfected with 1 µg of Vif expression vector or a control vector, plus 0.3 µg of A3F-231I-V5 expression vector encoding a V5-tagged A3F231I protein (lanes 1–6), or plus 0.3 µg of A3F231V-V5 expression vector encoding a V5-tagged A3F-231V protein (lanes 7–12). At 48 h after transfection, cell lysates were harvested for immunoblot analysis with indicated antibodies.

doi:10.1371/journal.pgen.1005921.g005

Influence of A3F 231V haplotype on HIV-1 Vif mediated degradation

HIV-1 Vif recruits Cul5-based CRL E3 ligase to target A3F for ubiquitination and degradation [37]. To determine whether A3F variants could alter A3F sensitivity to Vif mediated degradation, A3F I231-V5 and A3F V231-V5 expression vectors were co-transfected with several HIV-1 Vif expression vectors in HEK293T cells. A3F expression was determined by immunoblotting using anti-V5 antibodies. A3F I231-V5 and A3F V231-V5 were sensitive to strains 89.6, Yu2, and SG3 HIV-1 Vif mediated degradation but resistant to AE and BC HIV-1 Vif mediated degradation (Fig 5). Nevertheless, A3F V231-V5 was approximately 39% more resistant (based on normalized intensity) to HIV-1 AE Vif mediated degradation (Fig 5, lane 8) compared to A3F I231-V5 (Fig 5, lane 2). A3F V231-V5 was 30% more resistant to HIV-1 Yu2 Vif (Fig 5, lane 11) mediated degradation compared to A3F I231-V5 (Fig 5, lane 5). These data suggest that A3F variations at position 231 may influence the sensitivity of A3F to certain HIV-1 Vif mediated inactivation.

Influence of A3F 231V haplotype on HIV-1 hypermutation

The effect of rs2076101 on the level of HIV hypermutation was tested using paired human genotypes and viral sequences from 421 patients in the Swiss HIV Cohort Study. Counts of GA-to-AA hypermutations likely induced by A3F were quantified as in the Hypermut2 tool [38]. We observed 8.11 (out of 144.95 potential A3F specific G->A mutations, 5.5% hypermutation rate), 8.23 (5.6%) and 8.63 (5.3%) A3F specific G->A mutations in three rs2076101 genotype groups, respectively; the log odds ratio of hypermutation was not higher in individuals carrying the A3F 231V than those who don't (coefficient = 0.0047, $P = 0.88$, S2 Table).

Discussion

In vitro studies have identified A3F as one of the human APOBEC3 proteins with intrinsic anti-HIV-1 activity, and A3F unlike A3G is partially resistant to degradation by HIV-1 Vif [7,15–17]. In this population-based genetic study, we assessed the impact of A3F variation on HIV-1 disease progression in longitudinal HIV-1/AIDS natural disease cohorts. We found that the haplotype encoding A3F p. 231V and bearing 5 other variant alleles was significantly associated with slower progression to AIDS and lower viral load set-point. This protective effect is most pronounced for PCP, the most common AIDS-defining disease in the era before effective

antiretroviral therapy was available. Our results provide genetic evidence supporting an active role of human A3F in modulating HIV-1 disease *in vivo*. The observed population level associations point toward the physiological relevance of A3F, which has long been debated [7,15–21].

A3F suppresses HIV-1 replication through deaminase-induced G-to-A hypermutations in viral DNA, as well as deamination-independent impairment of viral reverse transcription and prevention of HIV DNA integration [15–17,39,40]. The mechanism by which the A3F variant haplotype influences AIDS pathogenesis is not clear. Mechanistically, the expectation would be that A3F variant protein isoforms have different abilities in inhibiting HIV replication, either by change of protein function or abundance. The association of p.231V with lower viral load supports the idea that A3F ancestral protein isoform containing p. 231V is more effective in restricting HIV replication compared to the derived protein isoforms. In an *in vitro* HIV-1 infectivity experiment, comparable or slight stronger inhibitory activity against an HIV-1 X4 strain with the wild type Vif was observed for A3F 231V (45% infectivity) compared with A3F 231I (57% infectivity) [21]. It is possible that over time the inhibitory effects of the variant protein may lead to detectable differences in disease outcomes. In a HIV-1 vif-A3F degradation assay, we have also observed that A3F I231V can influence A3F sensitivity to certain HIV-1 Vif proteins. Altered Vif sensitivity of A3F variation may be a contributing factor for the observed differences in HIV-1 viral loads in our study population. Direct Sanger sequencing of HIV-1 viral RNA in plasma of patients in the Swiss HIV cohort identified very low levels of G to A hypermutation, which may represent a small subset of the variation present in the integrated proviral HIV. Although HIV G to A variation mediated by A3F was not associated with the A3F I231V genotype, deeper sequencing provirus may be needed to accurately access the role of A3F genetic variants on HIV editing [41,42]. Further studies are also needed to determine if A3F variants affects A3F protein function in blocking reverse transcription and or HIV integration.

Pneumocystis pneumonia, or PCP, caused by the fungus *Pneumocystis jirovecii*, is the most common opportunistic infection in untreated AIDS patients [43,44]. PCP often occurs in those with low CD4+ cell counts $<200/\text{mm}^3$; in the MACS cohort, 30% of those patients developed PCP within 3 years after CD4+ cell counts dropping to $<200/\text{mm}^3$ [43]. A3F variation impact on PCP may be related with altered activity of A3F in the pulmonary inflammatory environment after CD4+ cell decline. The pathogenesis of PCP is marked with the immune-mediated pulmonary inflammation involving chemokines and cytokines [44]. Inflammatory stimuli such as lipopolysaccharide and interferon- α potently induce A3G and A3F protein expression [45]. Research will be needed to determine if A3F protein is activated in lung inflammation sites, affecting local viral replication and immune response.

Our study is limited in sample size, especially as the individual cohorts represent different HIV risk factors and transmission routes; and we are underpowered in the African American group to detect small or modest effect sizes. In African Americans, lack of significant association could be due to small sample size and lower PCP rate (11%) compared to European Americans (19%). It is also possible that other operational SNPs in African Americans may be tracking or interacting with A3F 231V. Our results require further replication in other well-powered cohorts with well-defined clinical AIDS outcome data. The strong linkage disequilibrium among multiple SNPs within the A3F gene with various potential functional effects makes it difficult to clearly identify the causal variants responsible for the phenotypic associations. In addition to the p. 231V site, other variant alleles carrying on the haplotype may also potentially affect A3F gene regulatory function as suggested by differential DNA-protein binding conferred by the intronic SNP rs5750728. Although A3F SNPs were not in strong LD with other A3 genes, the possibility of A3F SNPs tracking other SNPs in the A3 gene family region

cannot be completely ruled out due to the high homology among the gene family that pose technical difficulties in gene-specific genotyping, sequencing and assembly [46].

Our results may shed light on the relative contribution of A3 family members to HIV pathogenesis *in vivo*, a critical unanswered question. We and others have identified associations of a number of genetic variants in the A3 family genes with HIV-1 acquisition or disease progression [28–30,33]. The H186R of A3G was associated with rapid progression to AIDS in African Americans [30], higher viral load in South African infected women [47] and rapid disease progression in infected children [48]. The C40693T variant of A3G [28] [33] may be associated with increased risk to acquisition, whereas Haplotype I of A3H may confer resistance to infection [29]. A3B has been shown to restrict HIV-1 in experiments using 293 cells lines [3,7,49–51], but not in the more biologically relevant T-cell lines [52,53]. The impact of the A3B gene deletion on HIV-1 infection and progression is inconsistent, possibly due to differences in study designs and case-control selection. We previously reported that the homozygous A3B deletion (D/D) was associated with increased risk to HIV-1 acquisition by comparing HIV-1 exposed but uninfected individuals to HIV-1 incident seroconverters, and further that seroconverters progressed more rapidly to clinical outcomes [33]. In support of this, one study found that the D/D carriers had a six-fold greater likelihood of having lower CD4+ T cell levels [54]. Other studies found no associations for infection or progression, but these were suffered from frailty bias which might have inflated type 2 errors. For example, one study used healthy controls and long-term non-progressors but not fast progressors [55,56], and the other study used a control group of HIV-1 negative MSM who were nearly a decade younger than infected cases [57]; given equal exposure time, many of the control group may have become infected. This study did show tendency to lower CD4+ T cell depletion with the A3B deletion during 88 day follow-up [57]. The role of A3B and the A3B deletion in HIV disease warrants further clarification as their role in multiple virus, pathogens and cancers increasingly has been recognized [54,58–61].

In this study, we identified a common haplotype tagged by A3F 231V variant as a novel AIDS-modifying genetic factor in European Americans, which shows a similar trend in African Americans with HIV infection. A3G H186R is only found in individuals with African ancestry and is almost absent in European population [30], while the A3B gene deletion is more common in individuals from Asia and is less frequent in Europeans, and is nearly absent in populations with West African ancestry. The differences in APOBEC3 variant frequencies and haplotype structure among continental populations might have arisen from population-specific demographic events and or from viral selective pressure acting on the APOBEC3 genes. This study adds new evidence that A3F plays an important role in restricting pathogenesis of HIV-1 in its natural host. These data support the development of therapeutics targeting the Vif-A3F axis, along with current efforts on A3G-Vif.

Materials and Methods

Study participants

We used clinically well-characterized subjects from 6 pretreatment HIV/AIDS cohorts. Study participants (N = 4203) were enrolled in USA-based, prospective natural history HIV/AIDS cohorts: Multicenter AIDS Cohort Study (MACS) [62], the San Francisco City Clinic Cohort Study (SFCCC) [63], AIDS Link to the Intravenous Experience (ALIVE) [64], Hemophilia Growth and Development Study (HGDS) [65], the Multicenter Hemophilia Cohort Study (MHCS) [66] and DC gay (DCG) [67]. These cohorts were initiated during 1980's and participants were followed up semi-annually with blood draws for viral load and CD4+ T cell measurements and physical examinations at each visit. These treatment naïve cohorts have been used to identify multiple HIV-1/AIDS modifying genetic factors [30,68].

The study group includes 707 European American HIV-1 incident seroconverters, 281 African American incident seroconverters, 1135 HIV-uninfected at-risk individuals and 2076 HIV seroprevalent individuals who were HIV infected at study entry. The censoring date was the earliest of the date of the last recorded visit, or July 31, 1997 for the ALIVE cohort, or December 31, 1995 for all other cohorts, to avoid the confounding effect of highly active anti-retroviral therapy (HAART). A later censoring date was used for ALIVE cohort because few ALIVE participants received HAART prior to July 31, 1997 [69]. PCP prophylaxis status was only available for a subset of the MACS cohort; the first self-reported drugs used included Trimethoprim/sulfamethoxazole, aerosolized Pentamidine and Dapsone; the combination or switch-over of these drugs were used over the clinical course.

Ethics statement

Ethical approval for the study was obtained from the National Institute of Health Office of Human Subjects Research Protections with OHSRP #3314. Ethical approval was obtained from institutional review boards for each of the respective contributing centers in the International Collaboration for the Genomics of HIV. Written informed consent was obtained from all study participants.

Selection of *A3F* genetic variants

We selected six SNPs that were codon-changing, in a predicted splicing or regulatory site (based on SNPinfo web server [<http://snpinfo.niehs.nih.gov>]), considering SNP spacing and gene coverage. Several other nonsynonymous variants in *A3F* are reported in the 1000 genome project but were not genotyped here either because they were rare or tagged by other SNPs: rs2020390 (p. A108S) is in near-absolute LD ($r^2 = 0.94$) with rs2076101 (p.I231V) in Europeans. rs34182094 (p. A178T), rs12157816 (p. Y307C), rs13056825 (p.N346S) were absent or infrequent in Utah Europeans (CEU) and Yoruba Africans (YRI).

Genotyping of *A3F* genetic variants

SNPs were genotyped using the TaqMan allele discrimination assays on an ABI 7900HT sequencer detector system (Applied Biosystems), according to the manufacturer's protocol. For quality control, water controls were included on each plate and 10% of samples were duplicated. No water contamination or genotype mismatches between duplicates was observed. SNP rs2076101 was also genotyped by a second in-house designed TaqMan assay (sequences available upon request); the resulted genotypes were identical by two assays.

Electrophoretic mobility shift assays (EMSA)

EMSA was performed in nuclear extracts from HeLa cell stimulated with PMA using oligonucleotides carrying *A3F* rs5750728 T/C with an Infrared EMSA Kit (LI-COR), as previously described [70]. The double-stranded oligonucleotides used were (SNP allele underlined): 5'-CCCTTTGGCCAGTGCGTCCCACCACAT-3' and 5'-CCCTTTGGCCAGTGCGCCCACCACAT-3'.

Cell culture, transfection and immunoblot analysis

HEK293T (ATCC, catalog no. CRL-11268) cells were maintained in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) with 10% fetal bovine serum and penicillin/streptomycin (D-10 medium) and passaged upon confluence. DNA transfection in HEK293T cells was carried out using Lipofectamine 2000 (Invitrogen) as recommended by the manufacturer. For

degradation assays, HEK293T cells in 12-well plates were transfected with 1 μ g of Vif or 1 μ g of empty vector and 0.3 μ g of A3F expression vector at a 3:1 ratio. HEK293T cells were collected at 48 h after transfection. Cell lysates were lysed in 1x loading buffer (0.08 M Tris, pH 6.8, with 2.0% SDS, 10% glycerol, 0.1 M dithiothreitol and 0.2% bromophenol blue) and boiled for 5 min before separating the proteins by SDS-PAGE. Membranes were probed with various primary antibodies against proteins of interest. Secondary antibodies were alkaline phosphatase-conjugated anti-sheep, anti-rabbit or anti-mouse (Jackson ImmunoResearch) antibodies, and staining was carried out with 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT) solutions prepared from chemicals obtained from Sigma.

Detection of HIV-1 hypermutation

HIV-1 sequences covering the *pol* gene alone (in half of the patients) or the *pol* gene with additional other genes were obtained by the Sanger sequencing from plasma of patients in the Swiss HIV Cohort Study. A3F specific hypermutation was quantified as in the Los Alamos Hypermut2 tool [38]. Consensus nucleotide sequences of the samples and of the hxb2 reference sequence were translated to amino acid sequence from which a multiple alignment was created with muscle [71]. Using the amino acid sequence alignment we derived a codon-aware nucleotide alignment, which was used to count the number of possible hypermutation sites. We counted the number of trinucleotides matching the “GA[A or T or G]” pattern in the reference sequence as the number of potential A3F hypermutation sites. The number of actual A3F induced mutations in each sample was determined as a subset of potential A3F hypermutation sites where the first nucleotide was an A. It should be noted that GA \rightarrow AA editing can be mediated by A3D/F/H [39,41]. To quantify the background mutation level we used the following regular expression, which is the same as in the Hypermut2 tool: “G((CT)[ACGT])([AGC])”. The level of hypermutation of a patient sample was determined as an odds ratio: (number of A3F induced mutations/number of potential A3F specific sites) / (number of background mutations / number of potential background sites). Statistical tests were carried out on the logarithm of the latter quantity using linear regression. The regression included host PCA axes as covariates, the allele dosage of rs2076101, and an indicator variable indicating the SNP genotyping platform.

Statistical analysis

We performed analyses using SAS version 9.12 (SAS Institute, Cary, NC).

We assessed the influences of the A3F variants on disease progression to AIDS outcomes by the Cox proportional hazards model (Cox model) and Kaplan-Meier survival curve analyses. We only included seroconverters with known infection dates (midpoint estimation between last seronegative and first seropositive HIV test) in our analysis. The endpoint was clinical AIDS diagnosis using the Centers for Disease Control and Prevention (CDC) 1987 definition of AIDS [34], i.e. HIV-1 infection plus AIDS-defining illness or AIDS-related death. The median time from seroconversion to AIDS was 10 years in European American seroconverters. As the Kaplan-Meier survival analysis indicated that the haplotype tagged by rs2076101 best fitted a dominant genetic model, we compared the heterozygous (231I/V) and homozygous genotype (231V/V) state to the reference group of homozygotes (231I/I) in a Cox proportional hazards model. European American and African American groups were analyzed separately because the allele frequencies were different between the two groups. We included known genetic factors modifying AIDS progression as covariates in the adjusted Cox model analysis: *CCR5* Δ 32, *CCR5*-59029 (*CCR5*-2459, rs1799987), *HLA-B**27, *HLA-B**57, *HLA-B**35Px group (including *HLA-B**3502, *B**3503, *B**3504, and *B**5301), *HLA-C* [32] and *HLA* Class I

homozygosity for European American (reviewed in [72]; *HLA-B*57* and *HLA* Class I homozygosity for African American. We stratified the analyses by sex and by age at seroconversion: 0–20, >20–40, and > 40 years. *P* values for Cox model analysis were from the Wald test. Although our previous GWAS for association with HIV phenotypes in EA indicated minimal population substructure (genomic inflation factor $\lambda = 1.01$), we conservatively include the first two eigenvalues from a PCA analysis (EigenSoft) to correct for population stratification [73].

HIV-1 viral load set-point was defined as the mean \log_{10} -transformed copies of HIV-1 RNA in plasma measured between months 6–33 after seroconversion (2–5 measurements). Viral load measurements exceeding three-fold ($0.5 \log_{10}$) from the average of all remaining points were excluded, as previously suggested [35]. We used the *t* test for analysis of differences between viral load means of the group carrying 0 or 1 231V allele versus the 231I homozygote group.

The International Collaboration for the Genomics of HIV (ICGH) combined the GWAS data from 25 cohorts from Europe and North America, including cohorts used in this study [36]. The genotypes of the *A3F* SNPs from ICGH were obtained from the previous GWAS data [36]. The impact of the *A3F* SNPs on HIV-1 viral load in ICGH was assessed using a fixed-effect inverse-variance weighted meta-analysis as previously described [36]. A summary of the selected study groups and cohorts used in the meta-analysis is presented in S3 Table. Viral load data from a total of 10395 (7266 European Americans, 3129 African Americans) HIV-1 seropositives was used in the meta-analysis.

The impact of *A3F* variants on HIV-1 infection susceptibility was assessed by comparing allelic frequencies between the HIV-1 infected group comprising seroconverter and seroprevalent persons and the HIV-1 uninfected group composed of persons at risk for HIV. Odds ratios (OR) and *P* values were obtained by using a conditional logistic regression test. All *P* values were 2-tailed.

Supporting Information

S1 Table. Distribution of *A3F*-231I/V in the HIV-1 negative and positive groups.

(DOCX)

S2 Table. HIV-1 hypermutations detected from patients in the Swiss HIV Cohort Study, stratified by the *A3F*- rs2076101 genotypes.

(DOCX)

S3 Table. Summary of cohorts in the International Collaboration for the Genomics of HIV (ICGH) consortium used for HIV-1 viral load analysis.

(DOCX)

S1 Fig. Linkage disequilibrium in the *APOBEC3* gene region. Data was based on HapMap phase III data in the CEU (Utah Caucasian) population and was plotted with Haploview. The intensity of the box reflects the r^2 level and haplotype block was defined by 95% CI.

(TIFF)

S2 Fig. Frequency distribution of *A3F*-231I/V genotypes in European Americans of fast progressors and slow progressors.

(TIFF)

S3 Fig. Kaplan-Meier survival curves of *A3F*-231I/V genotype carriers for progression to PCP since HIV-1 infection in 278 African American seroconverters. RH and adjusted *P* values were obtained from the Cox proportional hazards model. *P* values for survival curves were obtained from the log-rank test.

(TIFF)

Acknowledgments

We would like to thank Paul J. McLaren (École Polytechnique Fédérale de Lausanne) for critical reading of the manuscript. We also thank other investigators in the IHAC consortium for sharing their phenotype-genotype association results.

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.

Author Contributions

Conceived and designed the experiments: PA CAW. Performed the experiments: PA EBR WZ. Analyzed the data: PA CWT IB. Contributed reagents/materials/analysis tools: SP JJG SD SB JF GDK XFY. Wrote the paper: PA CAW. Critically read, contributed to and commented on the manuscript: PA SP JJG SD GDK JF XFY CAW.

References

1. 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. PMID: [12167863](#)
2. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, et al. (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* 424: 99–103. PMID: [12808466](#)
3. Yu Q, Chen D, Konig R, Mariani R, Unutmaz D, et al. (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J Biol Chem* 279: 53379–53386. PMID: [15466872](#)
4. Suspene R, Guetard D, Henry M, Sommer P, Wain-Hobson S, et al. (2005) Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. *Proc Natl Acad Sci U S A* 102: 8321–8326. PMID: [15919829](#)
5. Chiu YL, Greene WC (2008) The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol* 26: 317–353. doi: [10.1146/annurev.immunol.26.021607.090350](#) PMID: [18304004](#)
6. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, et al. (2003) DNA deamination mediates innate immunity to retroviral infection. *Cell* 113: 803–809. PMID: [12809610](#)
7. Bishop KN, Holmes RK, Sheehy AM, Davidson NO, Cho SJ, et al. (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr Biol* 14: 1392–1396. PMID: [15296758](#)
8. Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, et al. (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424: 94–98. PMID: [12808465](#)
9. Conticello SG, Harris RS, Neuberger MS (2003) The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr Biol* 13: 2009–2013. PMID: [14614829](#)
10. Marin M, Rose KM, Kozak SL, Kabat D (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat Med* 9: 1398–1403. PMID: [14528301](#)
11. Sheehy AM, Gaddis NC, Malim MH (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med* 9: 1404–1407. PMID: [14528300](#)
12. Yu X, Yu Y, Liu B, Luo K, Kong W, et al. (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302: 1056–1060. PMID: [14564014](#)
13. Zhang W, Du J, Evans SL, Yu Y, Yu XF (2012) T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* 481: 376–379.
14. Sadler HA, Stenglein MD, Harris RS, Mansky LM (2010) APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis. *Journal of virology* 84: 7396–7404. doi: [10.1128/JVI.00056-10](#) PMID: [20463080](#)
15. 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. PMID: [15296757](#)
16. Wiegand HL, Doehle BP, Bogerd HP, Cullen BR (2004) A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *Embo J* 23: 2451–2458. PMID: [15152192](#)

17. Zheng YH, Irwin D, Kurosu T, Tokunaga K, Sata T, et al. (2004) Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J Virol* 78: 6073–6076. PMID: [15141007](#)
18. Gillick K, Pollpeter D, Phalora P, Kim EY, Wolinsky SM, et al. (2013) Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4(+) T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination. *Journal of virology* 87: 1508–1517. doi: [10.1128/JVI.02587-12](#) PMID: [23152537](#)
19. 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. *Journal of virology* 87: 444–453. doi: [10.1128/JVI.00676-12](#) PMID: [23097438](#)
20. Miyagi E, Brown CR, Opi S, Khan M, Goila-Gaur R, et al. (2010) Stably expressed APOBEC3F has negligible antiviral activity. *Journal of virology* 84: 11067–11075. doi: [10.1128/JVI.01249-10](#) PMID: [20702622](#)
21. Mulder LC, Ooms M, Majdak S, Smedresman J, Linscheid C, et al. (2010) Moderate influence of human APOBEC3F on HIV-1 replication in primary lymphocytes. *Journal of virology* 84: 9613–9617. doi: [10.1128/JVI.02630-09](#) PMID: [20592068](#)
22. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, et al. (2010) Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res* 38: 4274–4284. doi: [10.1093/nar/gkq174](#) PMID: [20308164](#)
23. Koning FA, Newman EN, Kim EY, Kunstman KJ, Wolinsky SM, et al. (2009) Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J Virol* 83: 9474–9485. doi: [10.1128/JVI.01089-09](#) PMID: [19587057](#)
24. Russell RA, Smith J, Barr R, Bhattacharyya D, Pathak VK (2009) Distinct domains within APOBEC3G and APOBEC3F interact with separate regions of human immunodeficiency virus type 1 Vif. *J Virol* 83: 1992–2003. doi: [10.1128/JVI.01621-08](#) PMID: [19036809](#)
25. Albin JS, Harris RS (2010) Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: implications for therapeutics. *Expert Rev Mol Med* 12: e4. doi: [10.1017/S1462399409001343](#) PMID: [20096141](#)
26. Sheehy AM, Erthal J (2012) APOBEC3 versus Retroviruses, Immunity versus Invasion: Clash of the Titans. *Molecular biology international* 2012: 974924. doi: [10.1155/2012/974924](#) PMID: [22720156](#)
27. Ross SR (2009) Are viruses inhibited by APOBEC3 molecules from their host species? *PLoS pathogens* 5: e1000347. doi: [10.1371/journal.ppat.1000347](#) PMID: [19390611](#)
28. Valcke HS, Bernard NF, Bruneau J, Alary M, Tsoukas CM, et al. (2006) APOBEC3G genetic variants and their association with risk of HIV infection in highly exposed Caucasians. *AIDS* 20: 1984–1986. PMID: [16988524](#)
29. Cagliani R, Riva S, Fumagalli M, Biasin M, Caputo SL, et al. (2011) A positively selected APOBEC3H haplotype is associated with natural resistance to HIV-1 infection. *Evolution* 65: 3311–3322. doi: [10.1111/j.1558-5646.2011.01368.x](#) PMID: [22023594](#)
30. An P, Bleiber G, Duggal P, Nelson G, May M, et al. (2004) APOBEC3G genetic variants and their influence on the progression to AIDS. *J Virol* 78: 11070–11076. PMID: [15452227](#)
31. An P, Winkler CA (2010) Host genes associated with HIV/AIDS: advances in gene discovery. *Trends Genet* 26: 13.
32. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, et al. (2007) A whole-genome association study of major determinants for host control of HIV-1. *Science* 317: 944–947. PMID: [17641165](#)
33. An P, Johnson R, Phair J, Kirk GD, Yu XF, et al. (2009) APOBEC3B Deletion and Risk of HIV-1 Acquisition. *J Infect Dis* 200: 1054–1058. doi: [10.1086/605644](#) PMID: [19698078](#)
34. CDC(1987) Revision of the CDC surveillance case definition for acquired immunodeficiency syndrome. *MMWR* 36 Suppl 1: 1S–15S.
35. Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, et al. (2009) Common genetic variation and the control of HIV-1 in humans. *PLoS Genet* 5: e1000791. doi: [10.1371/journal.pgen.1000791](#) PMID: [20041166](#)
36. McLaren PJ, Coulonges C, Ripke S, van den Berg L, Buchbinder S, et al. (2013) Association study of common genetic variants and HIV-1 acquisition in 6,300 infected cases and 7,200 controls. *PLoS Pathog* 9: e1003515. doi: [10.1371/journal.ppat.1003515](#) PMID: [23935489](#)
37. Liu B, Sarkis PT, Luo K, Yu Y, Yu XF (2005) Regulation of Apobec3F and human immunodeficiency virus type 1 Vif by Vif-Cul5-ElonB/C E3 ubiquitin ligase. *J Virol* 79: 9579–9587. PMID: [16014920](#)
38. Rose PP, Korber BT (2000) Detecting hypermutations in viral sequences with an emphasis on G→A hypermutation. *Bioinformatics* 16: 400–401. PMID: [10869039](#)

39. Kobayashi T, Koizumi Y, Takeuchi JS, Misawa N, Kimura Y, et al. (2014) Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. *J Virol* 88: 5881–5887. doi: [10.1128/JVI.00062-14](https://doi.org/10.1128/JVI.00062-14) PMID: [24623435](https://pubmed.ncbi.nlm.nih.gov/24623435/)
40. Mbisa JL, Bu W, Pathak VK (2010) APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. *J Virol* 84: 5250–5259. doi: [10.1128/JVI.02358-09](https://doi.org/10.1128/JVI.02358-09) PMID: [20219927](https://pubmed.ncbi.nlm.nih.gov/20219927/)
41. Cuevas JM, Geller R, Garijo R, Lopez-Aldeguer J, Sanjuan R (2015) Extremely High Mutation Rate of HIV-1 In Vivo. *PLoS Biol* 13: e1002251. doi: [10.1371/journal.pbio.1002251](https://doi.org/10.1371/journal.pbio.1002251) PMID: [26375597](https://pubmed.ncbi.nlm.nih.gov/26375597/)
42. Russell RA, Moore MD, Hu WS, Pathak VK (2009) APOBEC3G induces a hypermutation gradient: purifying selection at multiple steps during HIV-1 replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA. *Retrovirology* 6: 16. doi: [10.1186/1742-4690-6-16](https://doi.org/10.1186/1742-4690-6-16) PMID: [19216784](https://pubmed.ncbi.nlm.nih.gov/19216784/)
43. Phair J, Munoz A, Detels R, Kaslow R, Rinaldo C, et al. (1990) The risk of *Pneumocystis carinii* pneumonia among men infected with human immunodeficiency virus type 1. Multicenter AIDS Cohort Study Group. *N Engl J Med* 322: 161–165. PMID: [1967190](https://pubmed.ncbi.nlm.nih.gov/1967190/)
44. Thomas CF Jr., Limper AH (2004) *Pneumocystis pneumonia*. *N Engl J Med* 350: 2487–2498. PMID: [15190141](https://pubmed.ncbi.nlm.nih.gov/15190141/)
45. Mehta HV, Jones PH, Weiss JP, Okeoma CM (2012) IFN-alpha and lipopolysaccharide upregulate APOBEC3 mRNA through different signaling pathways. *J Immunol* 189: 4088–4103. doi: [10.4049/jimmunol.1200777](https://doi.org/10.4049/jimmunol.1200777) PMID: [22972924](https://pubmed.ncbi.nlm.nih.gov/22972924/)
46. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, et al. (2002) An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* 79: 285–296. PMID: [11863358](https://pubmed.ncbi.nlm.nih.gov/11863358/)
47. Reddy K, Winkler CA, Werner L, Misana K, Abdool Karim SS, et al. (2010) APOBEC3G expression is dysregulated in primary HIV-1 infection and polymorphic variants influence CD4+ T-cell counts and plasma viral load. *AIDS* 24: 195–204. doi: [10.1097/QAD.0b013e3283353bba](https://doi.org/10.1097/QAD.0b013e3283353bba) PMID: [19996938](https://pubmed.ncbi.nlm.nih.gov/19996938/)
48. Singh KK, Wang Y, Gray KP, Farhad M, Brummel S, et al. (2013) Genetic variants in the host restriction factor APOBEC3G are associated with HIV-1-related disease progression and central nervous system impairment in children. *Journal of acquired immune deficiency syndromes* 62: 197–203. doi: [10.1097/QAI.0b013e32827ab612](https://doi.org/10.1097/QAI.0b013e32827ab612) PMID: [23138837](https://pubmed.ncbi.nlm.nih.gov/23138837/)
49. Bogerd HP, Wiegand HL, Doehle BP, Cullen BR (2007) The intrinsic antiretroviral factor APOBEC3B contains two enzymatically active cytidine deaminase domains. *Virology* 364: 486–493. PMID: [17434555](https://pubmed.ncbi.nlm.nih.gov/17434555/)
50. Doehle BP, Schafer A, Cullen BR (2005) Human APOBEC3B is a potent inhibitor of HIV-1 infectivity and is resistant to HIV-1 Vif. *Virology* 339: 281–288. PMID: [15993456](https://pubmed.ncbi.nlm.nih.gov/15993456/)
51. Rose KM, Marin M, Kozak SL, Kabat D (2005) Regulated production and anti-HIV type 1 activities of cytidine deaminases APOBEC3B, 3F, and 3G. *AIDS Res Hum Retroviruses* 21: 611–619. PMID: [16060832](https://pubmed.ncbi.nlm.nih.gov/16060832/)
52. Hultquist JF, Lengyel JA, Refsland EW, LaRue RS, Lackey L, et al. (2011) Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J Virol* 85: 11220–11234. doi: [10.1128/JVI.05238-11](https://doi.org/10.1128/JVI.05238-11) PMID: [21835787](https://pubmed.ncbi.nlm.nih.gov/21835787/)
53. 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. doi: [10.1371/journal.ppat.1002800](https://doi.org/10.1371/journal.ppat.1002800) PMID: [22807680](https://pubmed.ncbi.nlm.nih.gov/22807680/)
54. Prasetyo AA, Sariyatun R, Reviono, Sari Y, Hudiyono, et al. (2015) The APOBEC3B deletion polymorphism is associated with prevalence of hepatitis B virus, hepatitis C virus, Torque Teno virus, and *Toxoplasma gondii* co-infection among HIV-infected individuals. *J Clin Virol* 70: 67–71. doi: [10.1016/j.jcv.2015.07.009](https://doi.org/10.1016/j.jcv.2015.07.009) PMID: [26305823](https://pubmed.ncbi.nlm.nih.gov/26305823/)
55. Itaya S, Nakajima T, Kaur G, Terunuma H, Ohtani H, et al. (2010) No evidence of an association between the APOBEC3B deletion polymorphism and susceptibility to HIV infection and AIDS in Japanese and Indian populations. *J Infect Dis* 202: 815–816; author reply 816–817. doi: [10.1086/655227](https://doi.org/10.1086/655227) PMID: [20684727](https://pubmed.ncbi.nlm.nih.gov/20684727/)
56. Winkler CA, An P, Buchbinder S, Donfield S, Goedert J, et al. (2010) Reply to Itaya et al. *Journal of Infectious Diseases* 202: 816–817.
57. Imahashi M, Izumi T, Watanabe D, Imamura J, Matsuoka K, et al. (2014) Lack of association between intact/deletion polymorphisms of the APOBEC3B gene and HIV-1 risk. *PLoS One* 9: e92861. doi: [10.1371/journal.pone.0092861](https://doi.org/10.1371/journal.pone.0092861) PMID: [24667791](https://pubmed.ncbi.nlm.nih.gov/24667791/)
58. Nik-Zainal S, Wedge DC, Alexandrov LB, Petljak M, Butler AP, et al. (2014) Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat Genet* 46: 487–491. doi: [10.1038/ng.2955](https://doi.org/10.1038/ng.2955) PMID: [24728294](https://pubmed.ncbi.nlm.nih.gov/24728294/)

59. Jha P, Sinha S, Kanchan K, Qidwai T, Narang A, et al. (2012) Deletion of the APOBEC3B gene strongly impacts susceptibility to falciparum malaria. *Infect Genet Evol* 12: 142–148. doi: [10.1016/j.meegid.2011.11.001](https://doi.org/10.1016/j.meegid.2011.11.001) PMID: [22108670](https://pubmed.ncbi.nlm.nih.gov/22108670/)
60. Zhang T, Cai J, Chang J, Yu D, Wu C, et al. (2013) Evidence of associations of APOBEC3B gene deletion with susceptibility to persistent HBV infection and hepatocellular carcinoma. *Hum Mol Genet* 22: 1262–1269. doi: [10.1093/hmg/dd513](https://doi.org/10.1093/hmg/dd513) PMID: [23213177](https://pubmed.ncbi.nlm.nih.gov/23213177/)
61. Burns MB, Temiz NA, Harris RS (2013) Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat Genet* 45: 977–983. doi: [10.1038/ng.2701](https://doi.org/10.1038/ng.2701) PMID: [23852168](https://pubmed.ncbi.nlm.nih.gov/23852168/)
62. Phair J, Jacobson L, Detels R, Rinaldo C, Saah A, et al. (1992) Acquired immune deficiency syndrome occurring within 5 years of infection with human immunodeficiency virus type-1: the Multicenter AIDS Cohort Study. *J Acquir Immune Defic Syndr* 5: 490–496. PMID: [1560346](https://pubmed.ncbi.nlm.nih.gov/1560346/)
63. Buchbinder SP, Katz MH, Hessel NA, O'Malley PM, Holmberg SD (1994) Long-term HIV-1 infection without immunologic progression. *AIDS* 8: 1123–1128. PMID: [7986410](https://pubmed.ncbi.nlm.nih.gov/7986410/)
64. Vlahov D, Graham N, Hoover D, Flynn C, Bartlett JG, et al. (1998) Prognostic indicators for AIDS and infectious disease death in HIV-infected injection drug users: plasma viral load and CD4+ cell count. *JAMA* 279: 35–40. PMID: [9424041](https://pubmed.ncbi.nlm.nih.gov/9424041/)
65. Hilgartner MW, Donfield SM, Willoughby A, Contant CF Jr., Evatt BL, et al. (1993) Hemophilia growth and development study. Design, methods, and entry data. *Am J Pediatr Hematol Oncol* 15: 208–218. PMID: [8498644](https://pubmed.ncbi.nlm.nih.gov/8498644/)
66. Goedert JJ, Kessler CM, Aledort LM, Biggar RJ, Andes WA, et al. (1989) A prospective study of human immunodeficiency virus type 1 infection and the development of AIDS in subjects with hemophilia. *N Engl J Med* 321: 1141–1148. PMID: [2477702](https://pubmed.ncbi.nlm.nih.gov/2477702/)
67. Goedert JJ, Biggar RJ, Melbye M, Mann DL, Wilson S, et al. (1987) Effect of T4 count and cofactors on the incidence of AIDS in homosexual men infected with human immunodeficiency virus. *JAMA* 257: 331–334. PMID: [3491911](https://pubmed.ncbi.nlm.nih.gov/3491911/)
68. An P, Duggal P, Wang LH, O'Brien SJ, Donfield S, et al. (2007) Polymorphisms of *CUL5* are associated with CD4+ T cell loss in HIV-1 infected individuals. *PLoS Genet* 3: e19. PMID: [17257057](https://pubmed.ncbi.nlm.nih.gov/17257057/)
69. Celentano DD, Galai N, Sethi AK, Shah NG, Strathdee SA, et al. (2001) Time to initiating highly active antiretroviral therapy among HIV-infected injection drug users. *AIDS* 15: 1707–1715. PMID: [11546947](https://pubmed.ncbi.nlm.nih.gov/11546947/)
70. An P, Goedert JJ, Donfield S, Buchbinder S, Kirk GD, et al. (2014) Regulatory Variation in HIV-1 Dependency Factor ZNRD1 Associates with Host Resistance to HIV-1 Acquisition. *The Journal of infectious diseases*.
71. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792–1797. PMID: [15034147](https://pubmed.ncbi.nlm.nih.gov/15034147/)
72. O'Brien SJ, Nelson GW (2004) Human genes that limit AIDS. *Nat Genet* 36: 565–574. PMID: [15167933](https://pubmed.ncbi.nlm.nih.gov/15167933/)
73. Troyer JL, Nelson GW, Lautenberger JA, Chinn L, McIntosh C, et al. (2011) Genome-wide association study implicates *PARD3B*-based AIDS restriction. *The Journal of infectious diseases* 203: 1491–1502. doi: [10.1093/infdis/jir046](https://doi.org/10.1093/infdis/jir046) PMID: [21502085](https://pubmed.ncbi.nlm.nih.gov/21502085/)
74. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 3.