**APOBEC3G expression is dysregulated in primary HIV-1 infection and polymorphic variants influence CD4⁺ T-cell counts and plasma viral load**

Kavidha Reddy, Cheryl A. Winkler, Lise Werner, Koleka Mlisana, Salim S. Abdool Karim, Thumbi Ndung'u, and the CAPRISA Acute Infection Study Team

**Objectives:** In the absence of HIV-1 virion infectivity factor (Vif), cellular cytosine deaminases such as apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) inhibit the virus by inducing hypermutations on viral DNA, among other mechanisms of action. We investigated the association of APOBEC3G mRNA levels and genetic variants on HIV-1 susceptibility, and early disease pathogenesis using viral load and CD4⁺ T-cell counts as outcomes.

**Methods:** Study participants were 250 South African women at high risk for HIV-1 subtype C infection. We used real-time PCR to measure the expression of APOBEC3G in HIV-negative and HIV-positive primary infection samples. APOBEC3G variants were identified by DNA re-sequencing and TaqMan genotyping.

**Results:** We found no correlation between APOBEC3G expression levels and plasma viral loads (r = 0.053, P = 0.596) or CD4⁺ T-cell counts (r = 0.030, P = 0.762) in 32 seroconverters. APOBEC3G expression levels were higher in HIV-negative individuals as compared with HIV-positive individuals (P < 0.0001), including matched pre and postinfection samples from the same individuals (n = 13, P < 0.0001). Twenty-four single nucleotide polymorphisms, including eight novel, were identified within APOBEC3G by re-sequencing and genotyping. The H186R mutation, a codon-changing variant in exon 4, and a 3' extragenic mutation (rs35228531) were associated with high viral loads (P = 0.0097 and P < 0.0001) and decreased CD4⁺ T-cell levels (P = 0.0081 and P < 0.0001), respectively.

**Conclusion:** These data suggest that APOBEC3G transcription is rapidly downregulated upon HIV-1 infection. During primary infection, APOBEC3G expression levels in peripheral blood mononuclear cells do not correlate with viral loads or CD4⁺ T-cell counts. Genetic variation of APOBEC3G may significantly affect early HIV-1 pathogenesis, although the mechanism remains unclear and warrants further investigation.

© 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins

AIDS 2010, 24:195–204

**Keywords:** apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G, HIV-1 subtype C, host proteins, mRNA expression, polymorphisms, primary infection

---

aHIV Pathogenesis Programme, Doris Duke Medical Research Institute, Nelson R. Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa. bLaboratory of Genomic Diversity, Science Applications International Cooperation-Frederick, National Cancer Institute-Frederick, Frederick, Maryland, USA, and cCentre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu-Natal, Durban, South Africa.

Correspondence to Thumbi Ndung'u, HIV Pathogenesis Programme, Doris Duke Medical Research Institute, Nelson R. Mandela School of Medicine, University of KwaZulu-Natal, Private Bag X7, Congella, Durban 4013, South Africa.

Tel: +27 31 260 4727; fax: +27 31 260 4036; e-mail: ndungu@ukzn.ac.za

Received: 17 December 2008; revised: 27 October 2009; accepted: 10 November 2009.

DOI:10.1097/QAD.0b013e3283353bba
Introduction

Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G), a human cytidine deaminase, has potent antiviral activity [1–4], and its polymorphic variants may modulate resistance to infection or disease progression among those infected [5]. In the absence of the HIV-1 accessory protein, virion infectivity factor (Vif), APOBEC3G, is packaged into budding virions, and subsequently deaminates dC to dU in the retroviral minus strand reverse transcripts in target cells. These substitutions register as dG (guanidine) to dA (adenine) transitions in retroviral plus stranded DNA [6]. Excessive G-to-A substitutions, known as hypermutation, are common among lentiviruses and introduce multiple termination codons across their genomes [6–9]. APOBEC3G and other cytidine deaminases may also inactivate lentiviruses by other mechanisms apart from hypermutation [10,11]. Hypermutated viral sequences have been identified in long-term nonprogressors and were predominant over time, suggesting that cytidine deaminases may play a role in viral control in vivo [12,13]. HIV-1 Vif counteracts APOBEC3G by blocking its encapsulation into virions, targeting the host protein to the ubiquitin pathway for proteasome-mediated degradation, resulting in the eradication of APOBEC3G and the loss of its anti-HIV activity [14–19].

Increased expression of APOBEC3G may overcome the effects of Vif by providing a competitive advantage over time and may eventually incapacitate the virus and suppress viremia [20]. Additionally, genetic variants of APOBEC3G may alter its function or level of expression, thereby enhancing or diminishing its anti-HIV activity [21–23].

Given that APOBEC3G is a key intrinsic antiretroviral host factor that possesses significant anti-HIV-1 activity in vitro, we reasoned that its antiviral effects in vivo might be particularly pronounced during the primary infection phase before adaptive immune responses become established. We therefore, investigated the hypothesis that high mRNA levels of APOBEC3G in peripheral blood mononuclear cells (PBMCs) of seroconverters are associated with low viral setpoint and high CD4+ T-cell counts during primary HIV-1 subtype C (HIV-1C) infection. We also investigated the effects of APOBEC3G genetic polymorphisms on HIV-1C pathogenesis in a South African cohort, in a population where the HIV-1 epidemic is severest.

Participants and methods

Study participants

The Centre for the AIDS Programme of Research In South Africa (CAPRISA) acute infection study [24] is an observational natural history study of HIV-1C infection established in Durban, South Africa, in 2004. Two hundred and forty-five women at high risk for HIV infection were enrolled into phase 1 of the study. Participants were screened monthly and seroconverters were identified by two HIV-1 rapid antibody tests, Determine (Abbott Laboratories, Tokyo, Japan) and Capillus (Trinity Biotech, Jamestown, New York, USA). Antibody-negative samples underwent pooled PCR testing for HIV-1 RNA (Ambiscreen v1.5; Roche Diagnostics, Rotkreuz, Switzerland). HIV-1 RNA-positive samples were subsequently confirmed by quantitative RNA (AmpliCor v2.0; Roche Diagnostics) and HIV enzyme immunoassay test (BEI 2000; Dade Behring, Marburg, Germany). Participants with acute HIV infection and those from other coincident cohorts were recruited into phase 2 on the basis of a reactive HIV antibody test within 3 months of a previously negative result or PCR positive in the absence of antibodies. The estimated time of seroconversion was determined as the midpoint between the last antibody-negative and first antibody-positive test or 14 days before the participant was PCR positive and antibody negative. Acutely infected participants are followed weekly for 3 weeks, fortnightly until 3 months after infection, monthly until 12 months after infection and thereafter quarterly for a maximum of 5.5 years. A flow diagram summarizing the study cohort and experiments is available in Fig. 1. Ethical approval was obtained from the University of KwaZulu-Natal’s Biomedical Research Ethics Committee and all participants provided written informed consent.

Sample collection, measurement of CD4 cell counts and plasma viral load

Blood was obtained by venipuncture and PBMCs were isolated by Ficoll–Histopaque (Sigma, St Louis, Missouri, USA) density gradient centrifugation and frozen until use. Viral load was determined using the automated COBAS AMPLICOR HIV-1 Monitor Test v1.5 (Roche Diagnostics). CD4 cells were enumerated by using the Multitest kit (CD4/CD3/CD8/CD45) on a four-parameter FACS Calibur flow cytometer (Becton Dickinson, Franklin Lakes, New Jersey, USA).

mRNA expression analysis

APOBEC3G mRNA expression was quantified in 30 HIV-negative participants and in longitudinal samples of 32 HIV-positive participants. Additionally, 13 of the 32 HIV-positive participants had preinfection (baseline) samples available. RNA was isolated from cryopreserved PBMCs immediately after thawing using Trizol reagent (Invitrogen, Carlsbad, California, USA) according to the manufacturer’s protocol. RNA was reverse transcribed to synthesize cDNA using the Quantitect Reverse Transcription Kit (Qiagen, Venlo, The Netherlands).

APOBEC3G mRNA expression was quantified by real-time PCR using SYBR Green chemistry (Roche Diagnostic) with gene expression assays and results normalized against glyceraldehyde-3-phosphate dehydrogenase (GAPDH).
Detection of APOBEC3G polymorphisms

A DNA panel of 30 HIV-positive and 30 HIV-negative samples was resequenced to identify single-nucleotide polymorphisms (SNPs) in APOBEC3G. Sequencing primers and protocols used were previously published [22]. Primers covered the putative 5' regulatory region, eight exons, exon-intron junctions, intron 1 and the 3' untranslated region of the APOBEC3G gene (GenBank sequences AL022318 and AL078641). These regions were amplified separately. Each 25 µl PCR reaction contained 10× PCR buffer, 1.5 or 2.5 mmol/l MgCl₂ (primer set dependent), 2.4 mmol/l deoxyribonucleoside triphosphate mix, 0.15 µl TaqGold and 5 µmol/l of each forward and reverse primer. This was amplified at 95°C for 10 min, 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 45 s and a final 10-min extension step at 72°C. PCR products were purified using exonuclease and shrimp alkaline phosphatase (Amersham Pharmacia, Uppsala, Sweden) and sequenced using overlapping primers and a BigDye Terminator Kit (Applied Biosystems, Foster City, California, USA). Sequencing primers, regions amplified...
and PCR conditions are available in supplementary material (Table S1).

**Genotyping of variants**

Five SNPs, identified by resequencing APOBEC3G, were further genotyped in 168 samples (37 HIV-positive and 131 HIV-negative). TaqMan SNP genotyping assays and PCR restriction fragment length polymorphism (RFLP) were used to determine genotypes for these polymorphisms. TaqMan assays were carried out according to the manufacturer’s protocol (Applied Biosystems). PCR primers, conditions and restriction enzymes used for the RFLP assay and details of TaqMan genotyping assays are available in supplementary material (Table S2). After preliminary statistical analysis, the H186R, rs6001417 and rs35228531 SNPs were selected for further genotyping on the basis of their strong association with viral load and CD4 cell count. H186R was genotyped in 250 samples (61 HIV-positive and 189 HIV-negative), rs6001417 was genotyped in 246 samples (58 HIV-positive and 188 HIV-negative) and rs35228531 was genotyped in 220 samples (52 HIV-positive and 168 HIV-negative).

**Statistical analysis**

APOBEC3G mRNA expression levels were compared between HIV-negative and HIV-positive individuals using a generalized estimating equation (GEE) model [25–27]. This analysis takes into account longitudinal (repeated) measures for each participant. The association between APOBEC3G mRNA levels and viral loads and CD4 cell counts was determined using rank correlation tests. Fisher’s exact test was used to test the association between HIV status and H186R (rs8177832), rs6001417 and rs35228531 genotypes. The genetic effects of these mutations on viral loads and CD4 cell counts were also determined by a GEE model taking into account longitudinal measures for each participant. The data were represented by a locally weighted scatterplot smoothing model, which was used to plot smooth curves over the data points. A Kaplan–Meier survival analysis was performed to assess the difference in CD4 decline between the H186R, rs6001417 and rs35228531 genotype groups, and Cox regression was used to acquire hazard ratios. All statistical analysis was performed using SAS version 9.1 (SAS Institute, Cary, North Carolina, USA) and graphs were generated using GraphPad Prism version 5.01 for Windows (GraphPad Software, San Diego, California, USA).

**Results**

**APOBEC3G expression**

There is paucity of data regarding the interplay between HIV-1 and APOBEC3G expression in vivo, particularly, during primary infection when rapid viral replication occurs, followed by resolution of viremia and establishment of steady-state equilibrium between the virus and the body’s immune responses. We therefore, investigated whether primary HIV-1C infection is associated with changes of APOBEC3G expression in PBMCs as compared with HIV-negative samples. Comparison of APOBEC3G mRNA levels between HIV-negative and HIV-positive individuals within 12 months of infection (primary infection) showed that APOBEC3G levels were significantly higher in HIV-negative individuals than in HIV-positive individuals (P < 0.0001) (Fig. 2a). Additionally, comparison of APOBEC3G expression levels in matched pre and postinfection samples of seroconverters also showed that APOBEC3G expression was significantly higher before seroconversion (P < 0.0001) (Fig. 2b). Further, there was no significant difference in APOBEC3G levels when compared between individuals who are persistently seronegative and preinfection samples of seroconverters (Fig. 2c). Comparison of APOBEC3G mRNA levels at various time points after infection (Fig. 2d) showed no significant change in expression levels over time.

There is conflicting data on the relationship between APOBEC3G mRNA levels in PBMCs versus plasma viral load and CD4 cell counts in chronic HIV-1 infection [20,28,29]. We thus next investigated whether there is a correlation between APOBEC3G mRNA levels and HIV-1 plasma viral load and CD4 cell counts during primary HIV-1 infection and found no association between these factors (data not shown).

**APOBEC3G variants**

APOBEC3G genetic variants have not been described in African populations. By resequencing and genotyping, we identified 24 SNPs within APOBEC3G in our cohort (Table 1). Sixteen of these SNPs were described previously and eight were novel. Further, An et al. [22] described seven SNPs within APOBEC3G in a USA-based study cohort, four of which were identified in our study cohort. Frequencies of these SNPs in our cohort were similar to those of the African–American group in the USA-based cohort (Table 1).

The codon-changing variant, H186R (rs8177832), in exon 4, had a frequency of 0.307 and was analyzed further, as the 186R allele was previously shown to have AIDS-accelerating effects [22]. Two other APOBEC3G variants, rs6001417 in intron 3 and rs35228531, an extragenic mutation located 3’ near the gene, having allele frequencies of 0.303 and 0.207, respectively, were also further analyzed after preliminary analysis indicated that they also had AIDS-accelerating effects.

**Effects of H186R, rs6001417 and rs35228531 mutations on primary HIV pathogenesis**

H186R genotypes were determined for 250 individuals (61 HIV-positive and 189 HIV-negative). We tested the association between HIV status and H186R genotypes...
Fig. 2. Comparison of APOBEC3G mRNA levels between (a) HIV-uninfected and HIV-infected participants, (b) preinfection and postinfection samples of HIV-infected individuals, (c) persistently seronegative individuals and preinfection samples of seroconverters and (d) longitudinal postinfection samples. In (b), the postinfection column represents postinfection samples from three time points for each individual. These time points are at 3, 6 and 12 months after infection. A generalized estimating equation analysis model was used to analyze these data, as this model takes into account longitudinal (repeated) measures for each participant.

and found no significant difference in the distribution of H186R genotypes between HIV-negative and HIV-positive individuals (P = 0.5838) (data not shown).

We also compared viral loads (Fig. 3a) and CD4+ T-cell counts (Fig. 3b) between genotypes. Viral load and CD4 measurements were classified into 0–3 months after infection and 3–12 months after infection time intervals to identify possible differences between genotypes during acute infection (0–3 months) and early chronic infection (3–12 months). There was an overall significant difference in viral loads between genotypes (P = 0.0097) across both time periods (Fig. 3a). In addition, during the first 3 months of infection, there was a significant difference in viral loads between the wild-type reference group (AA) and those homozygous for the mutation (GG) (P = 0.0362), with the GG genotype having higher viral loads than the AA genotype. There was also an overall significant difference in CD4 cell counts between the AA reference group and the GG genotypes (P = 0.0081) (Fig. 3b). Further, during the first 3 months of infection, there was a significant difference in CD4 cell counts between AA and GG genotypes (P = 0.0006), with GG genotype having lower CD4 cell counts than the AA genotype. The association with CD4+ T-cell count is consistent with observed genotype effects on viral load. At 3–12 months after infection, the GG genotype maintained its association with higher viral loads and lower CD4 cell counts when compared with the other genotypes. Additionally, the heterozygous AG genotype also displays significantly higher viral loads (P = 0.0005) and lower CD4 cell counts (P = 0.0078) when compared with the reference AA genotype at this later time period. Kaplan–Meier survival analysis (Fig. 3c) shows that those who have the GG genotype have a significantly shorter time to CD4 cell count less than 350 cells/μL. The hazard ratio for the GG group as compared with the AA or AG group is 3.84 [95% confidence interval (CI) 1.43–10.35, P = 0.0078].
<table>
<thead>
<tr>
<th>NCBI dbSNP ID</th>
<th>Location</th>
<th>Nucleotide change</th>
<th>Amino acid change</th>
<th>Hardy–Weinberg</th>
<th>Minor allele frequency</th>
<th>SNP</th>
<th>AA</th>
<th>EA</th>
<th>CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs5757463</td>
<td>5' Regulatory region</td>
<td>C/G/T</td>
<td></td>
<td>0.419</td>
<td>0.089</td>
<td>571G/C</td>
<td>0.091</td>
<td>0.063</td>
<td>0.116</td>
</tr>
<tr>
<td>rs7291971</td>
<td>5' Regulatory region</td>
<td>C/G</td>
<td></td>
<td>0.858</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6519166</td>
<td>5' Regulatory region</td>
<td>A/C</td>
<td></td>
<td>0.185</td>
<td>0.111</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs8142124</td>
<td>5' Regulatory region</td>
<td>C/T</td>
<td></td>
<td>0.574</td>
<td>0.246</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs5750743</td>
<td>5' Regulatory region</td>
<td>C/G</td>
<td></td>
<td>0.966</td>
<td>0.327</td>
<td>90C/G</td>
<td>0.319</td>
<td>0.340</td>
<td>0.263</td>
</tr>
<tr>
<td>rs8177322</td>
<td>Exon 4</td>
<td>A/G</td>
<td>H186R</td>
<td>0.040</td>
<td>0.307</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17496046</td>
<td>Exon 6</td>
<td>C/G</td>
<td>Q275E</td>
<td>0.028</td>
<td>0.169</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11545130</td>
<td>Exon 7</td>
<td>C/T</td>
<td>L371L</td>
<td>0.086</td>
<td>0.019</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs34100992</td>
<td>Intron 2</td>
<td>A/G</td>
<td></td>
<td>0.300</td>
<td>0.089</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6001417c</td>
<td>Intron 3</td>
<td>C/G</td>
<td></td>
<td>0.030</td>
<td>0.303</td>
<td>197193T/C</td>
<td>0.368</td>
<td>0.029</td>
<td>0.086</td>
</tr>
<tr>
<td>rs3736685</td>
<td>Intron 3</td>
<td>C/T</td>
<td></td>
<td>0.004</td>
<td>0.151</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1753780</td>
<td>Intron 6</td>
<td>A/G</td>
<td></td>
<td>0.112</td>
<td>0.454</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1753784</td>
<td>Intron 6</td>
<td>A/G</td>
<td></td>
<td>0.030</td>
<td>0.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1753788</td>
<td>Intron 7</td>
<td>A/G</td>
<td></td>
<td>0.722</td>
<td>0.031</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs35342554</td>
<td>3' near gene</td>
<td>C/T</td>
<td></td>
<td>0.155</td>
<td>0.290</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs35228531</td>
<td>3' near gene</td>
<td>C/T</td>
<td></td>
<td>0.519</td>
<td>0.207</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-ex1-snp1-REC</td>
<td>Intron 1</td>
<td>A/G</td>
<td></td>
<td>0.357</td>
<td>0.078</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-ex3-snp1-REC</td>
<td>Intron 4</td>
<td>A/G</td>
<td></td>
<td>0.060</td>
<td>0.164</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-ex6-snp1-REC</td>
<td>Intron 7</td>
<td>A/G</td>
<td></td>
<td>0.846</td>
<td>0.019</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-in13-snp3-REC</td>
<td>Intron 2</td>
<td>A/G</td>
<td></td>
<td>0.300</td>
<td>0.089</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-in13-snp1-Y</td>
<td>Intron 2</td>
<td>C/T</td>
<td></td>
<td>0.401</td>
<td>0.073</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-pm1-snp1-Y</td>
<td>5' regulatory region</td>
<td>C/T</td>
<td></td>
<td>0.344</td>
<td>0.082</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-3u-snp1-REC</td>
<td>Intron 7</td>
<td>A/G</td>
<td></td>
<td>0.581</td>
<td>0.048</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cem15-3u-snp2-W</td>
<td>Intron 7</td>
<td>A/T</td>
<td></td>
<td>0.518</td>
<td>0.056</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NCBI dbSNP ID, National Center for Biotechnology Information SNP database reference number; RFLP, restriction fragment length polymorphism; SNP, single-nucleotide polymorphism.

aComparisons with results from a published study.
bSNPs genotyped by RFLP assays.
cSNPs genotyped by TaqMan assays.
dT/C denotes complementary alleles.

Table 1. APOBEC3G single-nucleotide polymorphisms and their minor allele frequencies.

NCBI dbSNP ID, National Center for Biotechnology Information SNP database reference number; RFLP, restriction fragment length polymorphism; SNP, single-nucleotide polymorphism.

Minor allele frequency from the work of An et al. [22]a
Fig. 3. Comparison of viral loads and CD4⁺ T-cell counts between H186R (rs8177832) genotypes (a and b) and rs35228531 genotypes (d and e) over 0–12 months after infection using a locally weighted scatterplot smoothing model. Measurements were classified into 0–3 months after infection and 3–12 months after infection time intervals to identify differences between the genotypes during acute infection and early chronic infection. Confidence intervals are shown by the vertical lines and overall $P$ values are indicated. Kaplan–Meier survival curves were performed to assess the difference between the H186R (c) and rs35228531 (f) genotype groups in the event of a CD4⁺ T-cell count less than 350 cells/µl for more than two consecutive visits. Cox regression was used to acquire hazard ratios.
In addition, we tested the association of the rs6001417 mutation with HIV status, CD4 cell count and viral load (data not shown). This polymorphism was in near-perfect linkage disequilibrium with H186R, had no significant association with HIV status but the mutant genotype (GG) was associated with significantly higher viral loads and lower CD4 cell counts ($P < 0.0001$).

Analysis of the rs35228531 mutation indicated that individuals who are heterozygous for the mutation (CT) had significantly higher viral loads and lower CD4 levels over 12 months of infection as compared with the wild-type group (Fig. 3d and e). A Kaplan–Meier survival analysis (Fig. 3f) shows that those who have the CT genotype progress faster to a CD4 cell count below 350 cells/µl. The hazard ratio for the CT group as compared with the CC group is 2.28 (95% CI 0.85–6.07, $P = 0.1005$). Although this was not statistically significant, a trend is observable, as only 22.9% of the CC group progressed to a CD4 cell count below 350 cells/µl, whereas 47.1% of individuals with the CT genotype progressed to a CD4 cell count below 350 cells/µl.

**Discussion**

In this study, we were interested in describing the contribution of APOBEC3G to viral control during the critical primary infection phase, as well as understanding the kinetics of APOBEC3G expression before and after infection. Early HIV-1 infection has been associated with immobilization or profound dysregulation of antiviral immune responses [30,31]. Therefore, it is unclear whether HIV-1 infection would result in mobilization or dysregulation of intrinsic immunity factors such as APOBEC3G, which appear to be a key component of innate immunity. We hypothesized that at the critical phase of primary HIV-1 infection, APOBEC3G expression levels might correlate negatively with viremia and positively with CD4 T-cell counts.

Our data show that APOBEC3G mRNA levels are lower in HIV-positive primary infection PBMCs as compared with HIV-negative PBMCs. Furthermore, in matched pre and postinfection samples of seroconverters, APOBEC3G mRNA levels declined, suggesting a dysregulation of APOBEC3G mRNA during primary HIV-1 infection. Thus, our results are consistent with most of the earlier studies [20,28,29,32] showing that APOBEC3G expression levels were higher in HIV-uninfected individuals when compared with HIV-infected individuals, although contrary findings have also been reported [33]. Our study extends the findings from the earlier studies to the primary HIV-1 infection phase. We also show that in matched pre and postinfection samples, APOBEC3G levels declined, suggesting an active mechanism of APOBEC3G dysregulation rather than increased susceptibility to infection among those with low baseline (preinfection) levels of APOBEC3G. We found no significant differences in preinfection APOBEC3G mRNA levels between seroconverters and nonseroconverters followed longitudinally, suggesting that APOBEC3G mRNA levels per se are not associated with protection against HIV-1 infection.

Our results, therefore, lend support to previous observations that HIV-1 has specific mechanisms for countering APOBEC3G as a possible immune evasion strategy. We cannot also rule out the possibility that HIV-1 infection is associated with redistribution away from peripheral blood of cellular components that are enriched for APOBEC3G or that HIV-1 specifically targets such cells. Further studies will be needed to address how APOBEC levels change in different cell subsets following infection. Recent data from HIV-1-uninfected individuals reveal that APOBEC3G is broadly expressed in different hematopoietic subsets and that mRNA levels may not always be concordant with protein levels [34]. Thus future studies will need to address expression at both mRNA and protein levels in HIV-positive individuals.

Studies show that immune evasion against APOBEC3G is mediated by HIV-1 Vif, and that it is two-fold, involving translational and posttranslational inhibitory effects on APOBEC3G [17]. There have been no previous reports showing that HIV-1 inhibits APOBEC3G expression at the transcriptional level. Whether the APOBEC3G mRNA reduction seen in HIV-positive samples in this study is tied to or independent of ubiquitin proteasome degradation mediated via HIV-1 Vif may require further investigation, but our results suggest that a third mode of HIV-1 APOBEC3G inhibition may involve downregulation of mRNA expression. This observation is in contrast to in-vitro studies in which HIV-1 infection does not appear to affect APOBEC3G mRNA levels [17,35,36]. In agreement with two previous studies [28,29] performed in chronically infected patients, we found no correlation between APOBEC3G mRNA levels and viral load or CD4 cell counts in this primary infection cohort. In one study [20] that found a significant inverse correlation between APOBEC3G mRNA levels and viral load and a positive correlation with CD4 T-cell count, the investigators stimulated PBMCs with antibodies before RNA isolation, which may explain their contrary findings.

We also investigated the extent of genetic variation within APOBEC3G in a South African study cohort in which HIV-1C predominates. Sub-Saharan Africa is worst affected by the HIV-1 epidemic, and despite this situation, relatively few studies [5,22,37] have attempted to define genes that may affect HIV/AIDS outcomes in the local populations and yet their frequencies may vary according to ethnic background. In this cohort, we describe the frequencies of several SNPs and identify several novel SNPs that have not been described before.
Recently, it was reported that a genetic variant of \textit{APOBEC3G}, the \textit{HI186R} mutation, is associated with an AIDS-accelerating effect in African-Americans infected with HIV-1B [22]. This mutation, which occurred at a frequency of approximately 30\% in our HIV-1C-infected African cohort, was associated with significantly increased viral loads and decreased CD4\(^{+}\) T-cell counts, consistent with the earlier findings. The \textit{rs6001417} mutation is in linkage disequilibrium with \textit{HI186R} and, therefore, showed highly similar associations. The detrimental effects of this mutation are, therefore, observable during primary infection and become more prominent with progression to early chronic infection. It remains unclear how this polymorphism may affect the antiviral activity of \textit{APOBEC3G} [22,23,38,39], and our findings emphasize the need for further studies to address the possible mechanisms.

Our data, however, indicate that this SNP is out of Hardy–Weinberg equilibrium (HWE) \((P = 0.04)\) in our study population. We have eliminated genotyping error as a reason for this, as we have both resequencing and TaqMan genotyping confirmation. Further, genotypes that were obtained were consistent between duplicates and free of contamination, as negative controls did not amplify. When HWE was calculated separately for HIV-positive and HIV-negative individuals, we found that the HIV-positive group conformed to HWE \((P > 0.05)\), whereas individuals in the HIV-negative group were out of HWE \((P = 0.03)\). Therefore, the distortion seems to be due to the excess of the GG genotype in the negative group (26/31), which may suggest that the minor allele is protective against infection in homozygotes, although this was not significant in our cohort (odds ratio \(= 0.54, P = 0.23\)) using the AA as the reference group. This may be the reason for the distortion in HWE in the negative group in which we see an excess of GG genotypes. This suggests that this mutation may reduce susceptibility to HIV infection, but upon infection becomes detrimental and accelerates disease progression. A larger sample size will be required to resolve this issue.

Further, the \textit{rs35228531} mutation, which occurred at an allele frequency of approximately 20\% in our cohort, was associated with significantly high viral loads and low CD4 cell counts. This mutation, however, is extragenic, located \textit{3'} near the gene and may, therefore, affect transcriptional regulation or posttranscriptional modifications of this antiviral factor.

**Conclusion**

We have shown that HIV-1 infection is associated with rapid downregulation of \textit{APOBEC3G} expression at the transcriptional level. Studies to decipher the mechanisms involved and to possibly develop means for counteracting this are needed. During primary infection, \textit{APOBEC3G} expression levels in PBMCs do not correlate with viral loads or CD4\(^{+}\) T-cell counts. Importantly, this is the first study to describe \textit{APOBEC3G} genetic polymorphisms in an African setting, where HIV-1C prevalence and incidence rates are extremely high. This is the first study that indicates that \textit{APOBEC3G} may be an important HIV/AIDS-modifying gene during primary HIV-1 infection. Genetic variants of \textit{APOBEC3G} significantly affect early and late HIV-1 pathogenesis, although the mechanism remains unclear and warrants further investigation.

**Acknowledgements**

We thank Beth Binns-Roemer and Yuchun Zhou for excellent technical assistance. We thank the study participants, CAPRISA clinical and laboratory staff for providing specimens. The CAPRISA acute infection study was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH) and U.S. Department of Health and Human Services (grant #U19 AI 51794). K.R. was supported by the Columbia University–Southern African Fogarty AIDS International Training and Research Programme funded by the Fogarty International Center, NIH (grant #D43TW00231) and by a training grant from the East Coast Biotechnology Research Innovation Centre (LifeLab), funded by the South African Department of Science and Technology. T.N. is supported by the Hasso Plattner Foundation and holds the South African DST/NRF Research Chair in Systems Biology of HIV/AIDS. 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. 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. This Research was supported [in part] by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

K.R., T.N. and C.W. conceived the study, designed the experiments and interpreted the data. K.R. performed the experiments. K.R. and L.W. analyzed the data. K.R. and T.N. wrote the paper and all coauthors reviewed the manuscript. S.A.K., K.M. and members of the CAPRISA acute infection study team designed, established and maintained the study cohort and provided the samples.

**References**

### Table S1. PCR sequencing primers, regions amplified and PCR conditions.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Amplified Regions</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>T (°C)</th>
<th>Mg++</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEM15-Ex1</td>
<td>Exon 1, intron 1, exon 2, intron 2</td>
<td>gacggaatttcgctcttgtc</td>
<td>gggaagaaaggac.acacttg</td>
<td>60</td>
<td>1.5</td>
</tr>
<tr>
<td>CEM15-Ex4</td>
<td>Exon 3, intron 4, exon 4, intron 4</td>
<td>ggtgagagggggagttca</td>
<td>gaccccggctgggaacaggg</td>
<td>60</td>
<td>1.5</td>
</tr>
<tr>
<td>CEM15-Ex2</td>
<td>Intron 4, exon 5</td>
<td>cctcatgctctcttccttt</td>
<td>gtcgaacccaagaagcagt</td>
<td>60</td>
<td>1.5</td>
</tr>
<tr>
<td>CEM15-Ex6</td>
<td>Exon 5, intron 5, exon 6, intron 6</td>
<td>gctggaaggagaagcgaagacac</td>
<td>aagttctggtccgctgg</td>
<td>60</td>
<td>1.5</td>
</tr>
<tr>
<td>CEM15-Ex13</td>
<td>Intron 2</td>
<td>gccctggaaacccctctc</td>
<td>ttccctctcctcctgc</td>
<td>60</td>
<td>1.5</td>
</tr>
<tr>
<td>CEM15-Ex15</td>
<td>Intron 2, exon 3, intron 3</td>
<td>gccctgcggagaagcagacg</td>
<td>ttagcagcagcagcagag</td>
<td>60</td>
<td>1.5</td>
</tr>
<tr>
<td>CEM15-3U</td>
<td>Intron 7, exon 8, 3’ region</td>
<td>ccctcctgctgctctctctc</td>
<td>gcctcctgctgctctctctc</td>
<td>60</td>
<td>1.5</td>
</tr>
</tbody>
</table>

### Table S2. PCR primers, PCR conditions and restriction enzymes used for the RFLP assay, and details of TaqMan Genotyping Assays.

<table>
<thead>
<tr>
<th>NCBI dbSNP ID</th>
<th>Method</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>T (°C)</th>
<th>Mg++</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs5750743</td>
<td>RFLP</td>
<td>acgcctggtggccattactct</td>
<td>gcctctccttaaagtgaactc</td>
<td>62</td>
<td>2.0</td>
<td>AvaI</td>
</tr>
<tr>
<td>rs8177832</td>
<td>TaqMan Assay</td>
<td>C_2189646_10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17496046</td>
<td>TaqMan Assay</td>
<td>C_25649193_10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6001417</td>
<td>TaqMan Assay</td>
<td>C_30089175_10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3736685</td>
<td>TaqMan Assay</td>
<td>C_27489853_10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs35228531</td>
<td>TaqMan Assay</td>
<td>C_61215563_10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>