Presence of HIV-1 G-to-A mutations linked to APOBEC editing is more prevalent in non-B HIV-1 subtypes and is associated with lower HIV-1 reservoir

Basma Abdi, Sidonie Lambert-Niclot, Marc Wirden, Aude Jary, Elisa Teyssou, Sophie Sayon, Romain Palich, Roland Tubiana, Anne Simon, Marc-Antoine Valantin, et al.

Presence of HIV-1 G-to-A mutations linked to APOBEC editing is more prevalent in non-B HIV-1 subtypes and is associated with lower HIV-1 reservoir

Running head: Association of APOBEC3 footprints in HIV-1 DNA with non-B HIV-1 and low HIV-1 reservoir

Basma ABDI[1]*, Sidonie LAMBERT[2], Marc WIRDEN[1], Aude JARY[1], Elisa TEYSSOU[1], Sophie SAYON[1], Romain PALICH[3], Roland TUBIANA[3], Anne SIMON[4,] Marc-Antoine VALANTIN[3], Christine KATLAMA[3], Laurence MORAND-JOUBERT[2], Vincent CALVEZ[1], Anne-Geneviève MARCELIN[1], Cathia SOULIE[1]

[1]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP UMRS 1136), AP-HP, Hôpital Pitié Salpêtrière, Laboratoire de virologie, Paris, France;

[2]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP UMRS 1136), AP-HP, Hôpital Saint Antoine, Laboratoire de virologie, Paris, France;

[3]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, (IPLESP UMRS 1136), AP-HP, Hôpital Pitié Salpêtrière, Service des Maladies Infectieuses, Paris, France

[4]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, (IPLESP UMRS 1136), AP-HP, Hôpital Pitié Salpêtrière, Service de Médecine Interne, Paris, France.

Corresponding author : Dr Basma ABDI, Laboratoire de Virologie-CERVI, Hôpital Pitié Salpêtrière 45-83 Bd de l'hôpital 75013 Paris, France. Phone: 33 1 84 82 80 44. Fax: 33 1 42 17 74 11. Email: basma.abdi@aphp.fr

Word number: Abstract=250, Text=1781.

Abstract

Objectives

APOBEC3 editing activity contributes to sequences variation and viral diversification. We aimed to characterize virological and clinical factors associated with G-to-A mutations and stop codons in the HIV-1 reservoir, markers of APOBEC3 footprints, in order to better understand HIV-1 diversity among virologically suppressed HIV-1 infected patients.

Methods

Immuno-virological and clinical factors were compared between 92 patients harboring G-to-A mutations and stop codons (APOBEC+) in the reverse transcriptase gene and 92 patients without G-to-A mutations (APOBEC-) nor stop codons in their DNA genotypes.

Results

Patients were predominantly men (74.5%) and were mostly infected by B-subtype (69.0%), with 44.1% and 55.9% in APOBEC+ and APOBEC- groups, respectively. At time of HIV DNA genotypes, the total cell associated HIV-1 DNA load was 2.34 $\log_{10}$ copies/10^6 cells in median (IQR 1.85-2.67) and 33.2% of them had a detectable ultra-sensitive plasma viral load. Hypermutated sequences were identified in 28.2% in APOBEC+ group. The median of total cell-associated HIV-1 DNA level was significantly lower in APOBEC+ than APOBEC- group: 2.13 $\log_{10}$ copies/10^6 cells (IQR 1.60-2.60) *versus* 2.52 $\log_{10}$ copies/10^6 cells (IQR 2.19-2.71), (p<0.001). Presence of G-to-A mutations and stop codon was independently associated with HIV-1 subtype non B (p=0.017).

Conclusion

These results show an independent association between the presence of G-to-A mutations and stop codons with HIV-1 subtype non-B and low proviral DNA that could be explained by the APOBEC3 footprints and restriction of DNA synthesis and integration. However, further investigations are needed to study the contribution of Vif amino acid variability among HIV-1 subtypes.

Introduction

APOBEC3 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) is an important innate immune family of cytidine deaminase proteins. These cellular proteins block actively retroviral infection by hitchhiking newly produced viral particles.[1] Different APOBEC3 family members were distinguished but mainly two enzymes APOBEC3F and APOBEC3G that have slightly different substrate specificities, produce G-to-A transitions and inhibit profoundly the replication of HIV-1.[1,2] The antiviral activity of APOBEC3 is counteracted by the HIV viral infectivity factor (Vif) protein by mediating its degradation by the proteasome. When Vif is defective, the result is hypermutation, an inordinate number of identical transitions G to A.[1,3,4]

Several studies have shown that APOBEC3 contributes to sequences variation and viral diversification because of naturally occurring defective HIV variants.[3,5]

One study demonstrates also that defective genomes were systematically detected in 5 patients on long-term ART and this high level of defective genomes was correlated with a small size of HIV proviral DNA.[6]

A better understanding of HIV-1 diversity and the persistence of the HIV-1 reservoir is necessary to characterize the viruses in order to develop new therapeutic approaches and to better target HIV-1 patients susceptible to receive optimize eradication strategies. Thus, we aim here to compare patients with or without HIV-1 APOBEC mutations and stop codons in the HIV-1 reservoir in order to determine which virological or clinical factors could be associated with APOBEC editing activity.

Materials and Methods

Study design and patients

One hundred eighty-four HIV-1 patients were retrospectively studied at the Pitié-Salpêtrière hospital, Paris, France. The study was focused on the RT gene, the most region targeted by APOBEC editing activity according of the Stanford University list of signature APOBEC mutations (https://hivdb.stanford.edu/page/apobecs/). The only 5 G-to-A transitions identified in this list as RT DRAMs (D67N, E138K, M184I, E190G/S and M230) were used to select patients. Then, we included all HIV-1 patients between January 2011 and June 2019 with a Plasma Viral Load (pVL) ≤20 copies/mL under ART having at least one of these G-to-A drug resistance mutations (DRAMs) and stop codons in the reverse transcriptase gene (RT) in their HIV-1 DNA genotypic drug resistance test performed by Sanger sequencing according to the ANRS HIV Drug Resistance procedures (APOBEC+) (n=92). A second group of patients, with a pVL ≤20 copies/mL and with nor G-to-A DRAMs nor stop codons in their DNA genotypes were included adjusting on the genotype's time period (APOBEC-) (n=92).

Cumulative HIV-1 RNA and DNA DRAMs were interpreted using the latest ANRS resistance algorithm (http://www.hivfrenchresistance.org/).

This work was a retrospective non-interventional study with no addition to standard care procedures. The study was carried out in accordance with the Declaration of Helsinki.

Total HIV-DNA quantification

Cell-associated HIV-1 DNA was quantified by ultrasensitive real-time PCR (Generic HIV-DNA assay, Biocentric, Bandol, France) as previously described.[7]

Ultra-sensitive viral load

The pVL was quantified using the Cobas AmpliPrep/CobasTaqMan HIV-1 assay (Roche Diagnostics, Switzerland; lower detection limit of 20 copies/ml). Ultra-sensitive viral load (USpVL) in the range of 1–20 copies/ml was indicated qualitatively (presence or absence of detectable signal).[8]

Identification of hypermutated G-to-A sequences

Hypermutation was analyzed using the Hypermut Analysis and Detection of APOBEC-induced Hypermutation program (https://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html) in DNA sequences. A sequence was considered hypermutated if it registered a P value less than 0.05 on the Fisher's exact test that compared the number of G-to-A changes due to APOBEC3 versus HxB2 reference (K03455).

Statistical analysis

All reported values are medians with IQR for continuous variables and frequencies and percentages for categorical variables. We used the Mann–Whitney test and the Chi2 test for comparison between patients' groups APOBEC+ and APOBEC-. Univariable and multivariable models were used to identify factors associated with G-to-A mutations: age, sex, ethnicity, time from HIV diagnosis, duration of ART, time to undetectable pVL, zenith of pVL, Nadir of CD4 cells count, baseline CD4 cells count, baseline CD4/CD8 ratio, HIV-1 subtype and at time of HIV-1 DNA genotypic drug resistance test: HIV-1 DNA viral load, USpVL, CD4 cells count, CD4/CD8 ratio and hypermutation. Variable with an univariable p

value <0.2 were retained in the multivariable analysis. All reported p values are two-tailed, with significance set at 0.05. Analyses were performed using Statview for Windows (SAS Institute Inc., version 5.0.1, Cary, NC, USA).

Results

Patients' characteristics

One hundred eighty-four patients were included in the study. They were predominantly men (74.5%) and Caucasian (71.7%). They were HIV-1 diagnosed in median 22.8 years ago (15.5-27.2), ART treated for 19.1 years (13.7-22.2) and they had undetectable pVL for 6.9 years in median (3.1-11.0) (table 1). 183/184 (99.45%) of patients had received RT inhibitors in the past. The patients were mostly infected by B-subtype (69.0%), with 44.1% and 55.9% in APOBEC+ and APOBEC- (*p=0.017*), respectively. Details of frequency of HIV-1 subtypes are presented in table S1 in supplementary data. At time of HIV DNA genotype test, the total cell associated HIV-1 DNA load was 2.34 $\log_{10}$ copies/10^6 cells in median (IQR 1.85-2.67) and 33.2% of all patients had a detectable USpVL. The median of total cell-associated HIV-1 DNA level was significantly lower in APOBEC+ than APOBEC-: 2.13 (IQR 1.60-2.60) versus 2.52 $\log_{10}$ copies/10^6 cells (IQR 2.19-2.71), (*p<0.001*).

HIV-1 sequences characteristics

The distribution of HIV-1 G-to-A DRAMs and stop codons in the study participant's sequences of the APOBEC+ (n=92) was: 15.2% (14/92) D67N, 2.2% (2/92) E138K, 33.7% (31/92) M184I, 1.1% (1/92) E190G/S, 41.3% (38/92) M230I and 100% (92/92) stop codon. Stop codons were mostly found at the following amino acid positions: W71, W88, W153, W212, W229, W239 and W266.

The distribution of HIV-1 DRAMs not APOBEC related is shown in table S2 in supplementary data. Globally, there is no a significant difference between the two groups (APOBEC+ and APOBEC-) according to the DRAMs prevalence.

Hypermutated sequences were identified in 28.2% (26/92) in APOBEC+. As expected, no patients in APOBEC- had hypermuted HIV-1 sequences.

Risk factors associated with APOBEC mutations and stop codons

Univariate and multivariate logistic regression analyses were performed to assess independent associations between immune-virological and clinical data and APOBEC mutation's presence (table 2). Three factors were retained for the multivariable analysis: ethnicity (*p=0.018*), total cell-associated HIV-1 DNA level (*p=0.016*) and HIV-1 subtype (*p=0.016*). The multivariable analysis showed that G-to-A DRAMs and stop codons presence was independently associated with HIV-1 subtype and level of total cell associated HIV-1 DNA. The presence of G-to-A DRAMs and stop codons was 2.8 times higher (95% CI 1.25-6.35) for patients infected by HIV-1 non-B subtypes and inversely related to the level of total cell associated HIV-1 DNA (OR 0.34, 95% CI 0.18-0.63).

Discussion

In the context of viral diversity and the development of new therapeutic approaches for HIV-1 cure, careful consideration must be given to the identification and characterization of HIV-1 reservoir. The ability of APOBEC editing activity to introduce mutations into viral DNA may contribute to viral diversity. Here, we find in a large number of virologically well controlled HIV-1 patients that G-to-A mutations DRAMs and stop codons presence in RT sequences were associated with a smaller size of HIV-1 reservoir and HIV-1 non-B subtypes.

Our results are in line with other studies showing an inverse association between APOBEC3 protein activity and level of HIV proviral DNA. Fourati *et al.*, studied the relationship between the size of the reservoir and the frequency of defective genomes in 5 patients on successful ART and 5 untreated patients and concluded that a high level of defective genomes was correlated with a small size of HIV proviral DNA.[6] Other researchers quantified provirus and APOBEC3G levels in resting CD4+ T lymphocytes in HIV controllers and ART-suppressed non-controllers and concluded that the highest levels of APOBEC3G protein in resting memory CD4+ T cells were significantly associated with the lowest levels of DNA provirus.[9] These results could be explained by the innate APOBEC3 footprints on the viral genome. Indeed, APOBEC3 family is the host restriction factor that inhibits HIV-1 replication, DNA synthesis and integration by blocking viral plus-strand DNA transfer and inhibiting provirus establishment in the host genome.[10]

We showed also that the presence of HIV-1 G-to-A DRAMs linked to APOBEC editing were more prevalent in non-B HIV-1 subtypes. This finding could be explained by the genetic variability of Vif that counteracts the antiviral activity of APOBEC3 by proteasome degradation and polymorphism in APOBEC3 gene family members depending on the HIV-1 viral subtype. Previous works have focused on the contribution of Vif amino acid variability among HIV-1 subtypes and they all reported that Vif proteins derived from HIV-1 clinical and viral isolates of different subtypes varied in their activities against APOBEC3 but there were some discrepancies between results.[11–13] Mawuena *et al*. have found that several non-B subtypes Vif alleles have an efficient anti-APOBEC3 activity similar to the commonly used subtype B Vifs.[11] Other researchers showed that Vif derived from a subtype C molecular clone was less effective at overcoming APOBEC3-mediated inhibition than Vif derived from either subtype B or CRF02_AG molecular clones.[12] However, Yukie *et al*. found that Vif protein derived from subtype C strains harbored the most robust anti-APOBEC3 activity.[13]

Limitation of all these studies is the sample sizes which are not sufficiently representative of all the HIV-1 subtypes' diversity. In our study, we were not able to identify a specific non-B subtype that was associated with the presence of G-to-A DRAMs and stop codons. Thus, further investigations should be conducted focusing on particular non-B subtypes (i.e. C or CRF02_AG).

Our study had some limitations as the sample size, although all HIV-1 patients with APOBEC DRAMs and stop codons were selected on the studied period time. Multivariate models with relative small sample sizes tend to be unstable, thus the future study could benefit more from a larger sample size. Furthermore, the use of the classical Sanger sequencing technology could underestimate resistance and stop codons. The sensitivity of Sanger sequencing in detecting resistant variants within quasi-species of WT viruses is ~20%, whereas the sensitivity of the ultra-deep sequencing (UDS) assay allows the use of a sensitivity threshold as low as 1%.[14] Rodriguez and colleagues showed that UDS have better sensitivity in detecting mutations. They also found more stop codons by UDS 1% than bulk sequencing (44% versus 26%).[14] However, Sanger sequencing is a widely used technic in standard follow up of HIV-patients.

The ability of APOBEC3 to introduce G-to-A DRAMs and stop codons into HIV-1 viral DNA, and its contribution to viral diversity and viral pathogenesis is now clearly established. However, some factors could influence this contribution and this study showed that there is an independent association between the presence of G-to-A DRAMs and stop codons with HIV-1 subtype non-B and low proviral DNA. The biological mechanisms underlying these associations should be further investigated in order to better characterize HIV-1 reservoir and to help in developing new therapeutic approaches for HIV-1 cure.

Acknowledgments

Funding

This work has been funded by Agence Nationale de Recherche sur le Sida et les hépatites – Maladies Infectieuses Emergentes (ANRS-MIE, AC43).

Transparency Declaration

AGM and VC has received grants and honoraria from Janssen-Cilag, Gilead, MSD and VIIV Healthcare. LMJ has received honoraria from Janssen-Cilag, Gilead, MSD, Mylan and VIIV Healthcare. SLN has received grants from Gilead and VIIV Healthcare. RP has received travel grants and honoraria from Gilead, ViiV Healthcare, Janssen and Merck. All other authors: none to declare

References

1. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. *Nature Reviews Immunology* 2004; **4**: 868–77.

2. Fourati S, Lambert-Niclot S, Soulie C, *et al.* Differential impact of APOBEC3-driven mutagenesis on HIV evolution in diverse anatomical compartments. *AIDS* 2014; **28**: 487–91.

3. Simon V, Zennou V, Murray D, *et al*. Natural Variation in Vif: Differential Impact on APOBEC3G/3F and a Potential Role in HIV-1 Diversification. *PLOS Pathogens* 2005; **1**: e6.

4. Mehle A, Strack B, Ancuta P, *et al*. Vif Overcomes the Innate Antiviral Activity of APOBEC3G by Promoting Its Degradation in the Ubiquitin-Proteasome Pathway. *J Biol Chem* 2004; **279**: 7792–8.

5. Fourati S, Malet I, Lambert S, *et al.* E138K and M184I mutations in HIV-1 reverse transcriptase coemerge as a result of APOBEC3 editing in the absence of drug exposure. *AIDS* 2012; **26**: 1619–24.

6. Fourati S, Lambert-Niclot S, Soulie C, *et al.* HIV-1 genome is often defective in PBMCs and rectal tissues after long-term HAART as a result of APOBEC3 editing and correlates with the size of reservoirs. *J Antimicrob Chemother* 2012; **67**: 2323–6.

7. Avettand-Fènoël V, Chaix M-L, Blanche S, *et al.* LTR real-time PCR for HIV-1 DNA quantitation in blood cells for early diagnosis in infants born to seropositive mothers treated in HAART area (ANRS CO 01). *J Med Virol* 2009; **81**: 217–23.

8. Lambert-Niclot S, Grude M, Meynard J-L, *et al.* Ultrasensitive Human Immunodeficiency Virus Type 1 Viral Load as a Marker of Treatment Choice for Simplification Strategies. *Clin Infect Dis* 2018; **67**: 1883–9.

9. Pasquale MD, Kourteva Y, Allos T, *et al*. Lower HIV Provirus Levels Are Associated with More APOBEC3G Protein in Blood Resting Memory CD4+ T Lymphocytes of Controllers In Vivo. *PLOS ONE* 2013; **8**: e76002.

10. Mbisa JL, Bu W, Pathak VK. APOBEC3F and APOBEC3G Inhibit HIV-1 DNA Integration by Different Mechanisms. *J Virol* 2010; **84**: 5250–9.

11. Binka M, Ooms M, Steward M, *et al*. The Activity Spectrum of Vif from Multiple HIV-1 Subtypes against APOBEC3G, APOBEC3F, and APOBEC3H. *J Virol* 2012; **86**: 49–59.

12. Lisovsky I, Schader SM, Sloan RD, *et al.* HIV-1 Subtype Variability in Vif Derived from Molecular Clones Affects APOBEC3G-Mediated Host Restriction. *INT* 2013; **56**: 258–64.

13. Iwabu Y, Kinomoto M, Tatsumi M, *et al.* Differential Anti-APOBEC3G Activity of HIV-1 Vif Proteins Derived from Different Subtypes. *J Biol Chem* 2010; **285**: 35350–8.

14. Rodriguez C, Nere ML, Demontant V, *et al.* Ultra-deep sequencing improves the detection of drug resistance in cellular DNA from HIV-infected patients on ART with suppressed viraemia. *J Antimicrob Chemother* 2018; **73**: 3122–8.

	All patients n=184	APOBEC- group n=92	APOBEC+ group n=92	*P* value
Sex (Female versus Male), n (%)				
Female	47 (25.5)	23 (48.9)	24 (51.1)	0.866
Male	137 (74.5)	69 (50.4)	68 (49.6)	
Age (years), median (IQR)	58 (53-65)	59 (54-65)	56 (51-65)	0.063
Ethnicity (Caucasian versus Non Caucasian), n (%)				
Caucasian	132 (71.7)	66 (50)	66 (50)	1.000
Non Caucasian	52 (28.3)	26 (50)	26 (50)	
Time from HIV diagnosis (years), median (IQR)	22.8 (15.5-27.2)	24.6 (17.6-27.6)	21.7 (14.3-26.5)	0.072
Duration of ART (years), median (IQR)	19.1 (13.7-22.2)	19.78 (15.4-22.5)	18.7 (12.4-22.0)	0.187
Time to undetectable pVL (years), median (IQR)	6.95 (3.1-11.0)	6.77 (3.5-10.9)	7.0 (2.1-12.3)	0.845
Zenith of pVL (log_{10} copies/mL), median (IQR)	4.96 (4.25-5.44)	4.99 (4.25-5.38)	4.90 (4.25-5.47)	0.743
Nadir of CD4 (cells count /mm^3), median (IQR)	176 (78-289)	176 (61-281)	173 (80-297)	0.998
CD4 at baseline (cells count/mm^3), median (IQR)	315 (176-434)	333 (202-504)	312 (161-419)	0.401
CD4/CD8 ratio at baseline, median (IQR)	0.36 (0.20-0.60)	0.37 (0.20-0.60)	0.34 (0.20-0.60)	0.753
HIV-1 subtype, n (%)				
B	127 (69.0)	71 (55.9)	56 (44.1)	**0.017**
Non-B	57 (31.0)	21 (36.8)	36 (63.2)	
Total cell associated HIV-1 DNA (log_{10}/ 10^6cells), median (IQR)	2.34 (1.85-2.67)	2.52 (2.19-2.71)	2.13 (1.60-2.60)	**<0.001**
Ultra-sensitive pVL, n (%)	61 (33.2)	32 (52.5)	29 (47.5)	0.638

CD4 (cells count/mm^3), median (IQR)	617 (446-836)	669 (462-852)	602 (402-789)	0.092
CD4/CD8 ratio, median (IQR)	0.82 (0.58-1.26)	0.83 (0.60-1.41)	0.81 (0.54-1.18)	0.469

Table 1: Patient's characteristics. All reported values are medians with IQR for continuous variables and frequencies and percentages for categorical variables. Mann–Whitney and Chi2 tests are used for comparison between patients APOBEC+ and APOBEC-. Statistical significance set at 0.05 is indicated in bold. pVL= plasma Viral Load.

	Univariate analysis		Multivariate analysis	
	OR (95% IC)	*p*	OR (95% IC)	*p*
Sex (Female versus Male)				
Female	1			
Male	2.38 (0.62-9.20)	0.206		
Age (years)	0.98 (0.94-1.02)	0.464		
Ethnicity (Non Caucasian versus Caucasian)				
Non-Caucasian	1		1	
Caucasian	4.87 (1.31-18.08)	**0.018**	1.76 (0.76-4.05)	0.180
Time from HIV diagnosis (years)	1.00 (1.00-1.00)	0.825		
Duration of ART (years)	1.00 (1.00-1.00)	0.650		
Time to undetectable pVL (years)	1.00 (1.00-1.00)	0.928		
Zenith of pVL (log_{10} copies/mL)	0.94 (0.56-1.57)	0.819		
Nadir of CD4 (cells count /mm^3)	1.00 (0.99-1.00)	0.830		
CD4 (cells count/mm^3)	1.00 (0.99-1.00)	0.833		
CD4/CD8 ratio	1.87 (0.20-17.23)	0.578		
HIV sub-type (B versus Non_B)				
B	1		1	
Non_B	5.08 (1.36-18.98)	**0.016**	2.81 (1.25-6,35)	**0.013**
HIV parameters at time of HIV-1 DNA genotypic drug resistance				
Total cell associated HIV-1 DNA (log10/10^6cells)	0.30 (0.11-0.80)	**0.016**	0.34 (0.18-0.63)	**0.001**
Ultra-sensitive pVL	1.51 (0.55-4.17)	0.421		
CD4 (cells count/mm^3)	1.00 (0.99-1.00)	0.677		
CD4/CD8 ratio	0.77 (0.21-2.78)	0.695		

Table 2: Univariate and multivariate analysis to identify factors associated with APOBEC editing activity. Univariable P value <0.2 were retained in the multivariable analysis. Statistical significance set at 0.05 are indicated in bold. MSM= Men who have Sex with Men; pVL= plasma Viral Load

Supplementary materials

	All patients n=184	APOBEC- group n=92	APOBEC+ group n=92
B	127 (69)	71 (55.9)	56 (44.1)
CRF02_AG	25 (13.6)	11 (44)	14 (56)
A	3 (1.63)	1 (33.3)	2 (66.7)
D	3 (1.6)	1 (33.3)	2 (66.7)
F	1 (0.5)	1 (100)	0 (0)
G	4 (2.2)	1 (25)	3 (75)
H	1 (0.5)	0 (0)	1 (100)
CRF06_cpx	4 (2.2)	1 (25)	3 (75)
CRF13_cpx	2 (1.1)	1 (50)	1 (50)
CRF45_cpx	1 (0.5)	0 (0)	1 (100)
Non determined	13 (7.1)	4 (30.8)	9 ()

Table S1: **Frequency of HIV-1 subtype**. All reported values are frequencies and percentages for categorical variables

	All patients n=184	APOBEC- group n=92	APOBEC+ group n=92	*P* value
HIV-1 drug resistance associated mutations				
M41L	46 (25)	27 (58.7)	19 (41.3)	0.173
E44D	8 (4.3)	5 (62.5)	3 (37.5)	0.470
K65R	3 (1.6)	1 (33.3)	2 (66.7)	0.560
T69D	7 (3.8)	3 (42.9)	4 (57.1)	0.700
K70R	15 (8.2)	7 (46.7)	8 (53.3)	0.788
L74V	12 (6.5)	5 (41.7)	7 (58.3)	0.550
K101E	4 (2.2)	2 (50)	2 (50)	1.000
K103N	15 (8.2)	5 (33.3)	10 (66.7)	0.178
Y115F	2 (1.1)	1 (50)	1(50)	1.000
E138A	11 (6)	4 (36.4)	7 (63.6)	0.351
M184V	55 (29,9)	30 (54.5)	25 (45.5)	0.421
Y181C	10 (5.4)	5 (50)	5 (50)	1.000
Y188L	2 (1.1)	1 (50)	1 (50)	1.000
G190A	5 (2.7)	2 (40)	3 (60)	0.650
L210W	30 (16.3)	15 (50)	15 (50)	1.000
T215Y	44 (23.9)	24 (54.5)	20 (45.5)	0.489
T215D	6 (3.3)	2 (33.3)	4 (66.7)	0.406
T215E	2 (1.1)	1 (50)	1 (50)	1.000
T215H	3 (1.6)	1 (33.3)	2 (66.7)	0.560
T215S	19 (10.3)	6 (31.6)	13 (68.4)	0.090
K219Q	6 (3.3)	3 (50)	3 (50)	1.000
P225H	3 (1.6)	0 (0)	3 (100)	0.081

Table S2: Cumulative HIV-1 drug resistance associated mutations not APOBEC related in reverse transcriptase gene. All reported values are frequencies and percentages for categorical variables. Chi2 test is used for comparison between patients APOBEC+ and APOBEC-. Statistical significance set at 0.05 is indicated in bold.