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

Slim Fourati1–3*, Sidonie Lambert-Niclot1–3, Cathia Soulie1–3, Isabelle Malet1–3, Marc Antoine Valantin1,2,4, Benjamin Descours1,5,6, Zaina Ait-Arkoub1–3, Benoit Mory7, Guislaine Carcelain1,5,6, Christine Katlama1,2,4, Vincent Calvez1–3 and Anne Geneviève Marcelin1–3


*Corresponding author. Department of Virology, CERVI Pitie-Salpetriere Hospital, 83 Boulevard de l’Hôpital, 75013 Paris, France. Tel: +33142177418; Fax: +33142177411; E-mail: slim.fourati@psl.aphp.fr

Received 7 February 2012; returned 19 April 2012; revised 30 April 2012; accepted 10 May 2012

Objectives: Precise characterization of viruses present in reservoirs in long-term pretreated patients will be a major issue to consider in the context of viral eradication. We assessed the frequency of defective viruses present in cellular reservoirs.

Methods: Peripheral blood mononuclear cells (PBMCs) and rectal biopsy samples were compared between five patients on successful long-term highly active antiretroviral therapy (HAART) (>7 years without blips) and five untreated patients. Molecular cloning and sequencing of the reverse transcriptase region were used to detect the presence of and quantify in-frame stop codons in HIV quasi-species. The relationship between the size of the reservoir and the frequency of defective genomes was assessed.

Results: Defective genomes were systematically detected in all patients on long-term HAART in both compartments (PBMCs and rectal tissues), with a higher level of defective genomes per sample compared with PBMCs of untreated patients. A high level of defective genomes was correlated with a small size of HIV proviral DNA. Regarding the nucleotide context, guanine (G) to adenine (A) substitution at tryptophan positions was responsible for the appearance of 89% of all in-frame stop codons in HIV quasi-species. The relationship between the size of the reservoir and the frequency of defective genomes was assessed.

Conclusions: We propose a scenario whereby defective genomes accumulate during HAART treatment, eventually reaching a viral extinction threshold. In the context of viral eradication, measurement of the relative amounts of defective and non-defective viruses (by molecular cloning and ultradeep sequencing) should be used as a new criterion for eradicating HIV.

Keywords: viral extinction, quasispecies, viral compartmentalization

Introduction

HIV persists in reservoirs of latently infected CD4 T cells of patients receiving suppressive highly active antiretroviral therapy (HAART) and the possibility of eradicating these reservoirs is currently being investigated.1,2 One major issue in the future will be to define adequate markers to characterize the viruses that are still present in the reservoir after several years of antiretroviral treatment and/or after new therapeutic approaches that are being developed to cure HIV infection.

As HIV is an RNA virus, the quasispecies model predicts the existence of defective genomes in reservoirs. In this regard, measuring ‘non-viable’ proviral DNA should be of considerable value as a tool to characterize the size of reservoirs in terms of replication-competent genomes, most importantly after HAART treatment or in the context of viral eradication. Here, we investigate the
frequency of defective cellular reservoirs in clinical samples [peripheral blood mononuclear cells (PBMCs) and rectal tissue] and determine the relationship between the size of reservoirs and the frequency of defective genomes. We propose a model whereby effective HAART treatment leads to the accumulation of defective genomes.

**Methods**

**Patients and tissue collection**

Paired peripheral blood and rectal biopsy samples were analysed from the MUCCOR study (NCT01019044). Subjects were selected based on sample availability. All patients had signed a written consent, approved by the Agence Française de Sécurité Sanitaire des Produits de Santé (AFSSAPS), for blood and rectal biopsy collections. The Comité de Protection des Personnes (CPP) and AFSSAPS approved the study. As a control, we analysed peripheral blood cells recovered from five HIV-1-untreated patients (estimated duration of infection <2 years with median CD4 >200 cells/mm$^3$).

**Blood and rectal tissue processing and HIV-1 DNA quantification**

PBMCs were isolated by centrifugation on Ficoll Hypaque. Rectal biopsies were obtained all around the upper part of the rectum by anoscopy. Multiple biopsies were collected and were pooled for analysis. Rectal biopsies were digested and collected in complete RPMI 1640 medium supplemented with a cocktail of antibiotic and antimitotic agents. HIV-1 DNA was quantified on PBMCs and rectal biopsies as previously described.

**Proviral DNA isolation, cloning and sequencing**

The reverse transcriptase (RT) region was amplified using the Agence Nationale de Recherches sur le SIDA consensus technique and the PCR products were cloned as previously described. DNA clones were reamplified by PCR for sequencing (580 nt long-sequenced region). Overall, between 35 and 60 clones were investigated for each sample. Clonal diversity was evaluated by calculating the genetic distance among the representative sequences using Kimura’s two-parameter model taking into account transitional and transversional substitution rates (MEGA 4.0 software).

**Statistical analyses**

Quantitative data were compared between subgroups using the Mann–Whitney test. The correlation between HIV-1 DNA quantification and the rate of defective genomes was analysed using the non-parametric Spearman correlation coefficient and by regression of HIV-1 DNA quantification against the rate of defective genomes. Differences in G-to-A mutation frequencies were analysed using the Hypermut 2.0 program (http://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html).

**Results**

PBMCs and rectal biopsy samples recovered from five patients on successful long-term HAART with no detectable viraemia for at least 7 years (range 7–19 years) were analysed. As a control, we analysed PBMCs recovered from five HIV-1 untreated patients. To determine the level of defective HIV DNA per sample, we performed molecular cloning of a fragment of the RT region. We assessed the number of in-frame stop codons in each clone and the number of sequences harbouring in-frame stop codons in each sample. Overall, 578 clones were analysed. Comparisons between drug-naive patients and long-term-treated patients evidenced a systematic detection of in-frame stop codons in all patients on long-term HAART (5/5 patients) in two compartments (PBMCs and rectal tissue) with variable levels of in-frame stop codons per sample (median 21%; range 15%–100%), contrasting with only 1/5 PBMCs from untreated patients (median 0%; range 0%–15%) (P=0.007, Mann–Whitney test) (Figure 1a). In the long-term pretreated group, the frequency of defective genomes was similar between PBMCs and rectal tissue (P=0.6). In both compartments, some samples harboured a high number of stop codons: in one patient, PBMCs contained 90% of defective clones and in another patient 100% of clones from rectal tissue harboured five stop codons in the RT region per clone. Of note, stop codons were detected in bulk sequences in patients harbouring a great number of stop codons but were never detected in bulk sequences when the proportion of defective genomes was <36%. It is noteworthy that these results almost certainly underestimate the total proportion of defective genomes as the analysed region was limited to a fragment of the RT region.

Regarding the nucleotide context, one specific type of nucleotide substitution was largely preferential in inducing in-frame stop codons: the replacement of a guanine (G) with an adenine (A) at tryptophan positions (positions 71, 88, 153, 212 and 229 in RT) was responsible for the appearance of 89% of all in-frame stop codons observed in the viral sequences obtained from intra-patient independent clones and across patients’ samples. These substitutions occurred mostly (57%) in the context of G-to-A hypermutation (a mutational process in which G-to-A transitions far exceed all other mutations in a viral sequence) (P<0.05 using the Hypermut 2.0 program), thus reflecting the footprints of the host restriction factor APOBEC3 on the viral genome. Genomes harbouring stop codons were genetically more divergent than genomes without stop codons (Figure 1b, P<0.001, Mann–Whitney); this is consistent with previous data showing that G-to-A hypermutated sequences exhibit greater diversity than reference sequences. These results likely reflect the ability of APOBEC3 to contribute to HIV variability. Overall, these results highlight the important contribution of APOBEC3 editing as a major source of defective viruses in reservoirs in PBMCs and in rectal tissues.

We next examined the relationship between the frequency of defective HIV DNA and the size of the reservoir using a conventional HIV DNA quantification test. Overall, the median copy number of HIV proviral DNA for the study participants examined was 620 copies (range 36–5386) per 10$^6$ cells. There was an inverse correlation between the level of defective genomes and the size of HIV proviral DNA (P=0.035; $r^2=0.26$) (Figure 1c). These observations provide the first evidence of a relationship between a high proportion of defective proviruses in reservoirs and a small reservoir size in patients under HAART.

**Discussion**

In this study, we show that a great amount of the proviral DNA present after many years of antiretroviral treatment is defective
even in the major sites of HIV production [gut-associated lymphoid tissue (GALT)], likely reflecting APOBEC3 footprints on the viral genome. There is an apparent disconnection between the very high proportion of defective viral genomes observed in some patients with small reservoirs in this study and the fact that virological rebounds are usually observed when patients discontinue antiretrovirals. Possible explanations include the persistence of replication-competent viruses at low levels or in anatomical compartments not evaluated in our study (such as the ileum or the central nervous system). The HAART-treated patients in this study sustained undetectable levels of viraemia for a long period (>7 years) without any blip during the period of follow-up. In this context, there are few data evaluating virological rebound in case of treatment interruption. Of note, patients harbouring stop codons in bulk sequencing are patients that harboured the highest levels of defective viruses; if confirmed by other studies, bulk sequencing could be a simple marker that gives such information without the need to use quantitative measurements.

Our study also represents a new step towards understanding HIV dynamics in GALT. Overall, we observed that the amount of defective genomes was similar between rectal tissue and PBMCs. This is in line with other studies investigating HIV reservoirs in GALT tissues showing that HIV DNA in rectal cells is well correlated with HIV DNA in blood.\textsuperscript{3,7} Other studies observed that viral diversity does not significantly differ between GALT and blood during HAART.\textsuperscript{8,9} This suggests that cross-infection events occur between rectal tissue and PBMCs during HAART. Our study further underlines the contribution of a restriction factor, APOBEC3, acting in different compartments, thus

![Figure 1](image-url)
influencing viral dynamics and attenuation in patients under HAART. These results are consistent with the high level of APOBEC3G expression observed in different tissues, including rectal tissue.15

We propose a scenario whereby replicating viruses would be cleared over time on treatment, leaving an increasing representation of defective genomes. Before or at the beginning of any treatment, most of the viruses that are present in the body are competent for replication. During HAART, new cycles of infection are prevented or very limited;17 however, viral production may persist as HIV could re-emerge from the stochastic activation/proliferation of latently infected CD4 T cells leading to viral production and consequently to cell death. When the proviral genome is defective, generally as a consequence of APOBEC3 editing, long-lived infected cells persist longer than those infected by replication-competent viruses. Indeed, the transcription activity of G-to-A hypermutated HIV genomes is likely to be abolished during T cell activation because of a hypermutated non-functional long terminal repeat (LTR),12 thus avoiding an immune response (i.e. viral antigen presentation and consequent cytotoxic T lymphocyte killing). Additionally, specific mutations and stop codons may impede translation of viral messages. Following our hypothesis, after many years of HAART a decrease in the global number of HIV-infected cells harbouring proviral DNA would decrease concomitantly with an increase in the proportion of cells infected by defective viruses.

Our results may have significant clinical implications. New therapeutic approaches that are being developed to cure HIV infection should accelerate the accumulation of defective viruses in reservoirs, as they are expected to preferentially induce activation of HIV-infected CD4 cells. It is tempting to speculate that in reservoirs, as they are expected to preferentially induce activation of HIV-infected CD4 cells, it is possible that the remaining proviral DNA would be defective, reaching the viral extinction threshold (Figure 1d). This hypothesis mirrors an ancestral phenomenon of retroviral endogenization whereby APOBEC3 editing of endogenous retroviruses, such as human endogenous retrovirus HERV-K (HML-2), is thought to have contributed to fixation in germ-line cells millions of years ago.3,14 This study has its limitations. The sample size is small, impeding the statistical power. The inverse correlation found between the size of reservoir and defective genomes has to be confirmed in a larger population. Nevertheless, the findings here suggest that measurement of the proportion of defective viruses combined with total proviral DNA should be an interesting tool complementary to other markers (assessments of infectious virus in blood or the quantification of cell-associated HIV RNA) to estimate the size of ‘functional’ viruses in future clinical trials.

Acknowledgements
We thank ‘Objectif Recherche VACCin Sida’ (ORVACs), F. Lecardenon and A. Samri for their contribution in the MUCOVIR study. We thank G. Le Mallier and P. Grange for their technical assistance.

Funding
This work was supported by ‘Agence Nationale de recherche sur les SIDA et les hépatites virales’ (ANRS), the Association de Recherche en Virologie et Dermatologie (ARVD) and the European Community’s Seventh Framework Programme (FP7/2007–2013) under the project ‘Collaborative HIV and Anti-HIV Drug Resistance Network’ (CHAIN).

Transparency declarations
None to declare.

References
7 Avettand-Fenoel V, Prazuck T, Hocqueloux L et al. HIV-DNA in rectal cells is well correlated with HIV-DNA in blood in different groups of patients, including long-term non-progressors. AIDS 2008; 22: 1880–2.