

# HIV-1 viremia not suppressible by antiretroviral therapy can originate from large T cell clones producing infectious virus

Elias K. Halvas,<sup>1</sup> Kevin W. Joseph,<sup>1</sup> Leah D. Brandt,<sup>1</sup> Shuang Guo,<sup>2</sup> Michele D. Sobolewski,<sup>1</sup> Jana L. Jacobs,<sup>1</sup> Camille Tumiotto,<sup>1</sup> John K. Bui,<sup>3</sup> Joshua C. Cyktor,<sup>1</sup> Brandon F. Keele,<sup>4</sup> Gene D. Morse,<sup>5</sup> Michael J. Bale,<sup>6</sup> Wei Shao,<sup>7</sup> Mary F. Kearney,<sup>6</sup> John M. Coffin,<sup>8</sup> Jason W. Rausch,<sup>9</sup> Xiaolin Wu,<sup>2</sup> Stephen H. Hughes,<sup>6</sup> and John W. Mellors<sup>1</sup>

<sup>1</sup>Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. <sup>2</sup>Leidos Biomedical Research, Inc., Frederick, Maryland, USA. <sup>3</sup>New York-Presbyterian Hospital/Weill Cornell Medical Center, Weill Department of Medicine, New York, New York, USA. <sup>4</sup>AIDS and Cancer Virus Program, Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA. <sup>5</sup>NYS Center of Excellence in Bioinformatics and Life Sciences, Translational Pharmacology Research Core, University at Buffalo, Buffalo, New York, USA. <sup>6</sup>HIV Dynamics and Replication Program, National Cancer Institute, Frederick, Maryland, USA. <sup>7</sup>Advanced Biomedical Computing Science, Frederick National Laboratory for Cancer Research (FNLCR), Frederick, Maryland, USA. <sup>8</sup>Department of Molecular Biology and Microbiology, Tufts University, Boston, Massachusetts, USA. <sup>9</sup>Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, Maryland, USA.

**BACKGROUND.** HIV-1 viremia that is not suppressed by combination antiretroviral therapy (ART) is generally attributed to incomplete medication adherence and/or drug resistance. We evaluated individuals referred by clinicians for nonsuppressible viremia (plasma HIV-1 RNA above 40 copies/mL) despite reported adherence to ART and the absence of drug resistance to the current ART regimen.

**METHODS.** Samples were collected from at least 2 time points from 8 donors who had nonsuppressible viremia for more than 6 months. Single templates of HIV-1 RNA obtained from plasma and viral outgrowth of cultured cells and from proviral DNA were amplified by PCR and sequenced for evidence of clones of cells that produced infectious viruses. Clones were confirmed by host-proviral integration site analysis.

**RESULTS.** HIV-1 genomic RNA with identical sequences were identified in plasma samples from all 8 donors. The identical viral RNA sequences did not change over time and did not evolve resistance to the ART regimen. In 4 of the donors, viral RNA sequences obtained from plasma matched those sequences from viral outgrowth cultures, indicating that the viruses were replication competent. Integration sites for infectious proviruses from those 4 donors were mapped to the introns of the *MATR3*, *ZNF268*, *ZNF721/ABCA11P*, and *ABCA11P* genes. The sizes of the clones were estimated to be from 50 million to 350 million cells.

**CONCLUSION.** These findings show that clones of HIV-1-infected cells producing virus can cause failure of ART to suppress viremia. The mechanisms involved in clonal expansion and persistence need to be defined to effectively target viremia and the HIV-1 reservoir.

**FUNDING.** National Cancer Institute, NIH; Howard Hughes Medical Research Fellows Program, Howard Hughes Medical Institute; Bill and Melinda Gates Foundation; Office of AIDS Research; American Cancer Society; National Cancer Institute through a Leidos subcontract; National Institute for Allergy and Infectious Diseases, NIH, to the I4C Martin Delaney Collaboratory; University of Rochester Center for AIDS Research and University of Rochester HIV/AIDS Clinical Trials Unit.

## ► Related Commentary: p. 5665

**Conflict of interest:** JWM has served as a consultant to Gilead Sciences, Merck Co. Inc., Xi'an Yufan Biotechnologies, and Accelevir Diagnostics, and owns share options in Co-Crystal Pharmaceuticals, Inc and Abound Bio.

**Copyright:** © 2020, American Society for Clinical Investigation.

**Submitted:** March 13, 2020; **Accepted:** July 22, 2020;

**Published:** October 5, 2020.

**Reference information:** *J Clin Invest.* 2020;130(11):5847–5857.

<https://doi.org/10.1172/JCI138099>.

## Introduction

In most individuals, combination antiretroviral therapy (ART) suppresses HIV-1 viremia, measured as plasma HIV-1 RNA, to below the limit of detection of commercial assays (20–40 copies/mL), and confers major clinical benefits (1). Despite clinically effective ART, infected cells persist for the life of the individual and a small subset of the infected cells carry intact proviruses capable of producing infectious virus that can fuel viral rebound when ART is stopped (2–6). Recent work has shown that this reservoir of HIV-1

**Table 1. Demographics and clinical characteristics of donors evaluated for nonsuppressible viremia on antiretroviral therapy (ART)**

Donor ID	Age	Sex	Race	Date of HIV-1 diagnosis	Years on ART <sup>a</sup>	Years of detectable viremia <sup>b</sup>	Current ART regimen <sup>c</sup>
R-09	73	Male	White	01/2005	10	5.2	TDF/FTC/EFV
C-03	43	Male	White	06/1994	9	4.5	DRV/r/ETV/DTG
C-02	62	Male	African American	06/1990	18	7.4	TDF/FTC/EFV
F-07	59	Male	African American	06/1996	19	2.1	ABC/3TC/EFV
K-01	55	Female	African American	07/2011	2	1.3	TDF/FTC/DRV/r
P-08	45	Male	White	01/2005	10	2.8	TDF/FTC/EFV
T-05	59	Male	White	1992	17	5.1	TDF/FTC/ATV/r
A-06	59	Male	White	03/2003	12	3.6	TDF/FTC/ATV/r
<b>Median</b>	<b>59</b>				<b>11</b>	<b>4.1</b>	

<sup>a</sup>Years on ART at the time of initial evaluation. <sup>b</sup>Above the limit of detection of FDA-cleared plasma HIV-1 RNA assays. <sup>c</sup>ABC, abacavir; ATV/r, atazanavir/ritonavir; DRV/c, darunavir/cobicistat; DRV/r, darunavir/ritonavir; DTG, dolutegravir; EFV, efavirenz; ETV, etravirine; FTC, emtricitabine; TDF, tenofovir disoproxil fumarate; 3TC, lamivudine.

is sustained by the proliferation of clones of cells (7–11). Although most of the proviruses in the HIV-1 reservoir remain latent, a small fraction are transcriptionally active (12–15), leading to very low levels of viremia (usually about 1–3 HIV-1 RNA copies/mL of plasma) that can be detected only with sensitive research assays (16–18). We previously reported that one donor with advanced malignancy had clinically detected levels of virus in blood (>40 copies/mL) on ART that was composed of a mixture of drug-resistant virus and WT virus, the latter produced by a large clone of cells carrying an intact, infectious provirus (10). Similar instances have not been published.

Current guidelines for treatment of HIV-1 recommend sustained suppression of viremia to below the limit of detection of FDA-approved plasma HIV-1 RNA assays (<20–40 copies/mL depending upon the assay). HIV-1 care providers become concerned when viremia is above this level, particularly when it is sustained. Medication review, tests for the presence of HIV-1 drug resistance, and adherence counseling are the standard management approaches used in such cases. Some providers may change or intensify ART regimens by adding additional antiretrovirals in an attempt to achieve greater viremia suppression, although this approach has been ineffective in decreasing low-level viremia in several clinical trials (19–21). The current study arose from requests by local HIV-1 care providers to investigate the cause of persistent viremia (detectable >6 months at >20 copies/mL) above the limit of detection of FDA-approved HIV-1 RNA assays in individuals who reported consistent adherence to their medication. We found that the nonsuppressible viremia could originate from large clones of infected cells rather than ongoing viral replication as a consequence of medication nonadherence or HIV-1 drug resistance.

## Results

We evaluated 8 individuals with nonsuppressible viremia, plasma HIV-1 RNA above 20 copies/mL and detectable for more than 6 months, despite reported adherence to ART and absence of drug resistance to the current regimen. The characteristics of the 8 study participants are summarized in Table 1. All individuals were on long-term suppressive ART (median 11 years, range 2–19 years) before exhibiting nonsuppressible viremia for more than 6

months (median 4.1 years, range 1.3–7.4 years). Median pre-ART plasma HIV-1 RNA was 172,506 copies per mL (range 30,375–16,700,000) and median nadir CD4<sup>+</sup> T cell count was 212 cells/mm<sup>3</sup> (range 10–314 cells/mm<sup>3</sup>), consistent with substantial immunodeficiency at the time of ART initiation. Other virologic and immunologic characteristics of the donors are provided in Table 2. At the time of referral, the median plasma HIV-1 RNA was 87 copies/mL (range 43–197 copies/mL), despite 1 or 2 prior regimen changes in 5 of the 8 donors during the duration of the nonsuppressible viremia. Median HIV-1 DNA and cell-associated HIV-1 RNA levels per million PBMCs were 1,458 copies (range 373–2,505 copies) and 261 copies (range 29–1,162 copies), respectively, quantified by qPCR targeting the 3' end of integrase (22). The median infectious units per million (IUPM) CD4<sup>+</sup> T cells was 0.5 (range 0.1–18.1 IUPM) (23, 24). The median average pairwise distances (APDs) of all proviral sequences was 1.95% (range 0.5%–2.4%), consistent with diverse HIV-1 populations arising after long-term infection (Supplemental Table 1; supplemental material available online with this article; <https://doi.org/10.1172/JCI138099DS1>). Immunophenotyping of PBMCs showed no major differences between donors and healthy HIV-1-negative controls (25–32) in the frequencies of T cells, B cells, or NK cells or in expression levels of surface activation markers (CD25, CD69, CD38, and CD107a), with the exception of higher HLA-DR expression (Supplemental Table 3), which is consistent with prior observations in HIV-1-positive individuals on effective ART (33).

*Identical plasma-derived HIV-1 sequences contribute to detectable viremia.* All 8 donors were evaluated using a number of different approaches (Supplemental Figure 1, A–F, and refs. 23, 24, 34–37). In all 8 donors, groups of identical HIV-1 RNA *gag-pro-pol* sequences derived from single HIV-1 RNA templates (single-genome sequencing [SGS]) (Supplemental Figure 1) were found in plasma samples, suggesting, but not proving, that the identical viruses were derived from clones of infected cells (Figures 1, 2, 3, and 4, Supplemental Figures 2–5, Table 3, and ref. 35). The identical RNA sequences comprised a substantial proportion of the total viral sequences in the plasma in all donors (median 58.6%, range 37.5%–100%) (Table 3). Longitudinal plasma sampling showed no changes in the identical viral sequences over a median sampling duration of 0.8 years

**Table 2. Immunologic and virologic characteristics of donors referred for nonsuppressible viremia**

Donor ID	Pre-ART Plasma HIV-1 RNA (cps/mL)	Nadir CD4 <sup>+</sup> T Cells (cells/mm <sup>3</sup> )	Current CD4 <sup>+</sup> T Cells (cells/mm <sup>3</sup> )	Plasma HIV-1 RNA at Referral (cps/mL) <sup>A</sup>	HIV-1 DNA (cps/10 <sup>6</sup> PBMCs) <sup>B</sup>	Cell-Associated HIV-1 RNA (cps/10 <sup>6</sup> PBMCs) <sup>B</sup>	IUPM <sup>C</sup>
R-09	97,000	105	380	197	1,533	139	18.1
C-03	16,700,000	10	416	62	2,505	1,162	1.4
C-02	30,375	286	1,022	184	373	29	0.1
F-07	117,068	314	1,023	52	1,603	1,112	3.8
K-01	147,189	133	533	68	1,383	74	0.6
P-08	604,000	172	444	106	1,056	382	0.4
T-05	197,826	299	1,105	113	650	109	0.4
A-06	1,877,100	251	831	43	1,825	630	0.2
<b>Median</b>	<b>172,506</b>	<b>212</b>	<b>682</b>	<b>87</b>	<b>1,458</b>	<b>261</b>	<b>0.5</b>

<sup>A</sup>Plasma HIV-1 RNA copies/mL at time of referral determined by FDA-approved Roche CAP/CTM v2.0 or Abbott M2000 platforms. <sup>B</sup>HIV-1 DNA and cell-associated RNA copies/million peripheral blood mononuclear cells (PBMCs) measured by quantitative polymerase chain reaction (22). <sup>C</sup>Infectious units per million (IUPM) total CD4<sup>+</sup> T cells by quantitative viral outgrowth assay (qVOA) (23, 24).

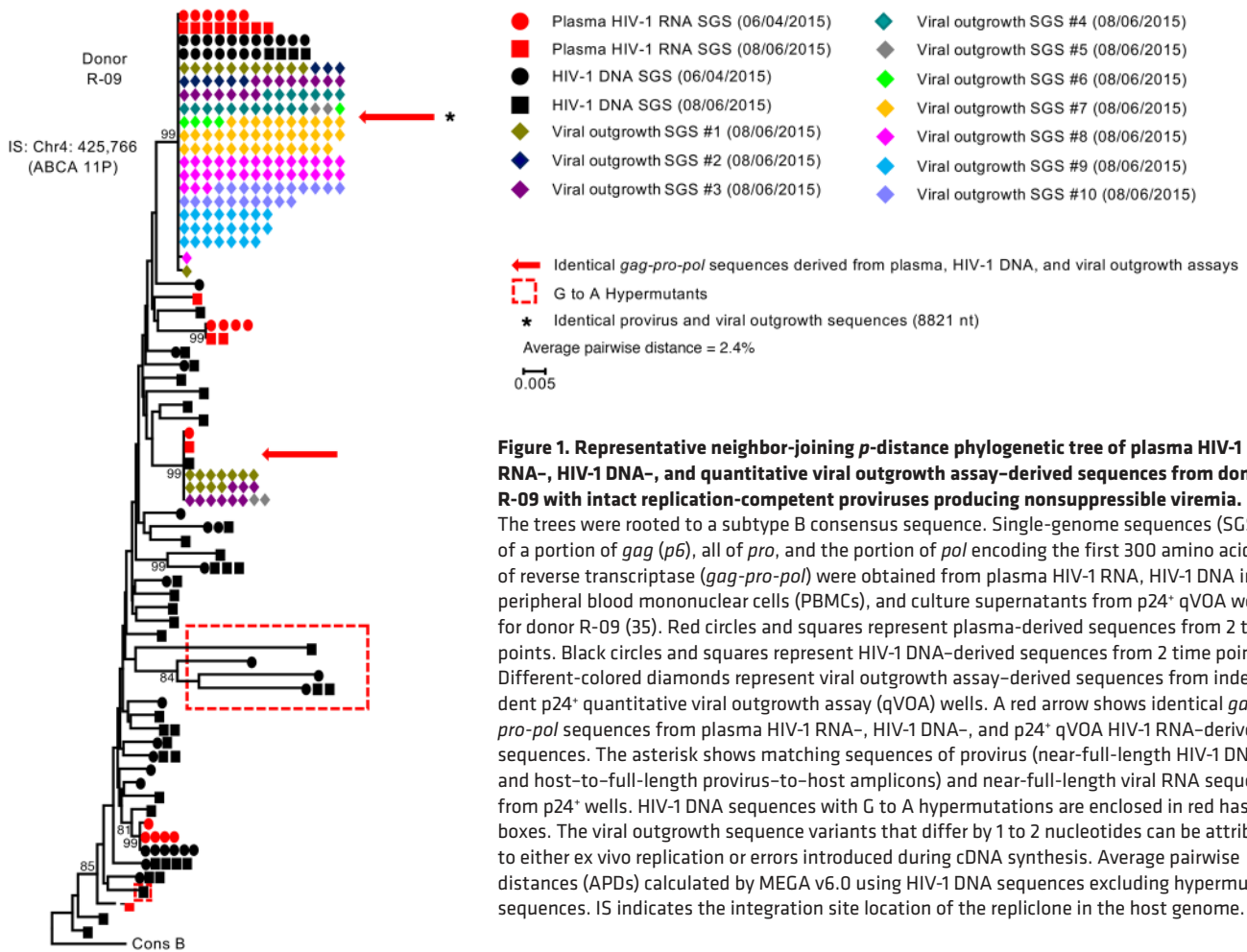
(range 0.2–1.2 years). Drug susceptibilities of the plasma viruses, predicted by sequence analysis, showed no evidence of drug resistance to each donor's current ART regimen, except for a group of identical plasma-derived sequences from donor F-07 (Supplemental Table 1 and ref. 38). These sequences from donor F-07 predicted low-level resistance (D67N/K70R/K219Q) to only abacavir, comprising 53.8% of the total plasma sequences (Table 3 and Supplemental Table 1). Drug concentrations for all donors were within the expected therapeutic range based on target trough concentrations (Supplemental Table 2 and refs. 39, 40).

*Identical plasma-, proviral-, and viral outgrowth-derived HIV-1 sequences.* Genomic DNA (gDNA) from PBMCs or purified CD4<sup>+</sup> T cells and HIV-1 RNA from p24<sup>+</sup> quantitative viral outgrowth assay (qVOA) wells were analyzed by *gag-pro-pol* SGS (Supplemental Figure 1, B–D). Matching identical *gag-pro-pol* sequences were found in proviral DNA and plasma HIV-1 RNA in 6 of the 8 donors (K-01, C-02, C-03, F-07, P-08, and R-09) (Figures 1–4 and Supplemental Figures 2 and 3). In 3 of these donors (C-02, C-03, and R-09), the identical viral sequences in plasma also matched multiple qVOA-derived sequences from p24<sup>+</sup> wells (Table 3 and Figures 1–3). In 1 donor (F-07), there were identical proviral and qVOA-derived sequences, but no sequences were found in plasma that matched (0 of 39). A different provirus from this donor matched plasma viral sequences, but these sequences were not found in p24<sup>+</sup> qVOA wells (Table 3 and Figure 4).

*Persistent viremia from intact, replication-competent HIV-1 proviruses.* Near-full-length (NFL) viral sequence analyses were performed on samples from the 4 donors (R-09, C-03, C-02, and F-07) for which there were identical *gag-pro-pol* sequence matches between plasma and p24<sup>+</sup> qVOA cultures. Phylogenetic analyses of sequences from donor R-09 suggested strongly that the persistent viremia was being produced by a clone of cells that carried an intact, replication-competent provirus. Six NFL HIV-1 DNA and 4 NFL HIV-1 RNA sequences from p24<sup>+</sup> qVOA wells were identical over 8,821 nt and matched the *gag-pro-pol* sequences found in plasma in longitudinal samples obtained 2 months apart (Figure 1 and Table 3). Similarly, phylogenetic evidence for a clonal origin of persistent viremia was also obtained for donors C-03

and C-02 from whom identical NFL proviral and qVOA-derived sequences (8,818 and 8,906 nt, respectively) were found that matched plasma-derived *gag-pro-pol* sequences (Figures 2 and 3 and Table 3). Identical viral RNA sequences were obtained from the plasma of these 2 donors over periods of 14 and 10 months, respectively. Full-length SGS of genomic RNA in plasma was not performed because of the technical challenge of low HIV-1 RNA template number (62–197 copies/mL) and the requirement for multiple sequences. However, SGS was performed for full-length HIV-1 *integrase* (donor C-03) and *envelope* (donors C-02, C-03, and R-09) (2.6 kb) from the same longitudinal samples and these sequences matched the full-length sequences of these proviruses (data not shown). In donor F-07, one NFL proviral sequence (8,819 nt) matched identical *gag-pro-pol* plasma-derived sequences from 2 longitudinal samples taken 13 months apart, but matching viral sequences were not found in p24<sup>+</sup> qVOA wells (Figure 4 and Table 3). Identical NFL sequences (8,804 nt) of a different provirus were found in blood that matched the qVOA-derived *gag-pro-pol* sequences, but as noted above, no matches were found in the viral RNA in plasma. Collectively, these analyses suggest that there are clones of cells in the blood of these 4 donors that can carry replication-competent proviruses, which we have termed replicones, and that in 3 of the donors the clones produce enough virions to cause clinically detected nonsuppressible viremia. Coreceptor tropism analyses (Geno2Pheno) of *env* sequences derived from NFL proviral or viral NFL RNA sequences from qVOAs indicated that viruses were CCR5-tropic (Supplemental Table 1 and ref. 41).

*Nonsuppressible viremia can originate from clones of infected cells.* We performed integration site analysis (ISA) (Supplemental Figure 1E) on genomic DNA extracted from purified CD4<sup>+</sup> T cells from the donors (R-09, C-03, C-02, and F-07) with plasma and/or proviral sequences that matched viruses obtained from qVOAs using a multiple-displacement amplification (MDA) approach (34, 42, 43). In samples from each of these 4 donors, we found multiple identical integration sites for the proviruses whose sequences matched the identical sequences in plasma and qVOA. Table 4 and Supplemental Figure 6 show the locations and orientations of proviruses in the host genome (44, 45). All integration



**Figure 1. Representative neighbor-joining *p*-distance phylogenetic tree of plasma HIV-1 RNA-, HIV-1 DNA-, and quantitative viral outgrowth assay-derived sequences from donor R-09 with intact replication-competent proviruses producing nonsuppressible viremia.** The trees were rooted to a subtype B consensus sequence. Single-genome sequences (SGSs) of a portion of *gag* (*p6*), all of *pro*, and the portion of *pol* encoding the first 300 amino acids of reverse transcriptase (*gag-pro-pol*) were obtained from plasma HIV-1 RNA, HIV-1 DNA in peripheral blood mononuclear cells (PBMCs), and culture supernatants from p24<sup>+</sup> qVOA wells for donor R-09 (35). Red circles and squares represent plasma-derived sequences from 2 time points. Black circles and squares represent HIV-1 DNA-derived sequences from 2 time points. Different-colored diamonds represent viral outgrowth assay-derived sequences from independent p24<sup>+</sup> quantitative viral outgrowth assay (qVOA) wells. A red arrow shows identical *gag-pro-pol* sequences from plasma HIV-1 RNA-, HIV-1 DNA-, and p24<sup>+</sup> qVOA HIV-1 RNA-derived sequences. The asterisk shows matching sequences of provirus (near-full-length HIV-1 DNA and host-to-full-length provirus-to-host amplicons) and near-full-length viral RNA sequences from p24<sup>+</sup> wells. HIV-1 DNA sequences with G to A hypermutations are enclosed in red hashed boxes. The viral outgrowth sequence variants that differ by 1 to 2 nucleotides can be attributed to either *ex vivo* replication or errors introduced during cDNA synthesis. Average pairwise distances (APDs) calculated by MEGA v6.0 using HIV-1 DNA sequences excluding hypermutated sequences. IS indicates the integration site location of the replicone in the host genome.

sites identified were within introns of known host genes including *ABCA11P*, *ZNF268*, *MATR3*, and *ZNF721* for donors R-09, C-03, C-02, and F-07, respectively. The exact location and full-length sequence of each of the intact proviruses was confirmed by amplifying and sequencing the entire provirus and flanking host sequences directly from genomic DNA extracted from PBMCs (Supplemental Figure 1F). We determined the frequency of these integrated intact proviruses compared with all other proviruses (Table 4) and found them to be a minor fraction of all infected cells (0.03%–1.1%) (34). We also estimated the total size of each of the replicones within each of the 4 individuals based on the approximate sizes of the CD4<sup>+</sup> T cell populations in each of the donors and the frequency of the specific, intact provirus per million CD4<sup>+</sup> T cells (Table 2 and refs. 22, 46). The total size of the replicones ranged from approximately 50 million to 350 million CD4<sup>+</sup> T cells (Table 4). These findings indicate that individual replicones are large enough to make an important contribution to the HIV-1 reservoir even though they comprise a very small fraction of all infected cells.

**Discussion**

Here we describe what we believe is a new cause of HIV-1 plasma viremia observed in standard clinical practice that is not suppressible by ART, and we show that the viremia can arise from large

clones of HIV-1-infected cells. Extensive virologic analyses of the 8 individuals referred for nonsuppressible viremia on ART revealed important insights for patient management and for efforts to cure HIV-1 infection. Although each of the 8 individuals have unique features, the consistent finding in all donors was that nonsuppressible plasma viremia consisted of one or more large groups of identical viral sequences. The largest group of identical viral sequences in plasma comprised 37.5% to 100% (median 58.6%) of all HIV-1 RNA in the plasma (Table 3). Similar findings of identical viral sequences in plasma were previously reported among individuals on effective ART, although the origin of the identical viruses was not identified (47, 48). In the current study, drug resistance to the donors' ART regimen was not evident by sequence analysis nor was medication nonadherence, since antiretroviral drug concentrations measured in random plasma samples were within the therapeutic range (Supplemental Tables 1 and 2). Longitudinal analyses of the viral sequences in all donors showed no evidence of virus evolution over time, indicating that these viruses originated from a stable, nonevolving reservoir of infected cells. For all 3 donors with infectious clonal viremia, identical sequence matches were found between plasma HIV-1 RNA, proviral DNA, and HIV-1 RNA in outgrowth cultures across the 1.5 kb *p6/protease/reverse transcriptase* amplicon and the full-length 2.6 kb *envelope* amplicon.



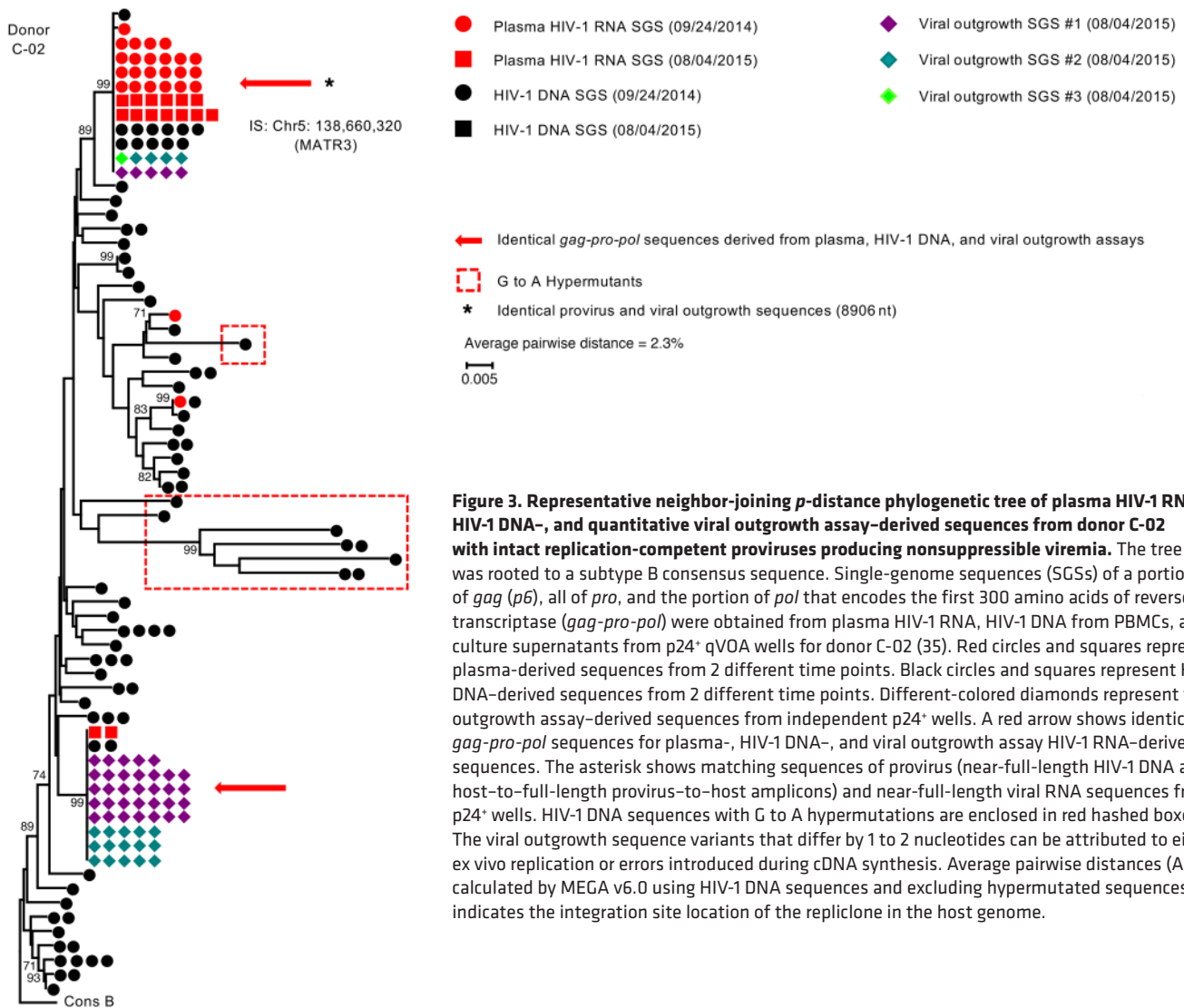
**Figure 2. Representative neighbor-joining *p*-distance phylogenetic tree of plasma HIV-1 RNA-, HIV-1 DNA-, and quantitative viral outgrowth assay-derived sequences from donor C-03 with intact replication-competent proviruses producing nonsuppressible viremia.**

The tree was rooted to a subtype B consensus sequence. Single-genome sequences (SGSs) of a portion of *gag* (*g6*), all of *pro*, and the portion of *pol* encoding the first 300 amino acids of reverse transcriptase (*gag-pro-pol*) were obtained from plasma HIV-1 RNA, HIV-1 DNA in PBMCs, and culture supernatants from p24<sup>+</sup> qVOA wells for donor C-03 (35). Red circles and squares represent plasma-derived sequences from 2 time points. Black circles and squares represent HIV-1 DNA-derived sequences from 2 time points. Different-colored diamonds represent viral outgrowth assay-derived sequences from independent p24<sup>+</sup> quantitative viral outgrowth assay (qVOA) wells. A red arrow shows identical *gag-pro-pol* sequences from plasma HIV-1 RNA-, HIV-1 DNA-, and p24<sup>+</sup> qVOA HIV-1 RNA-derived sequences. The asterisk shows matching sequences of provirus (near-full-length HIV-1 DNA and host-to-full-length provirus-to-host amplicons) and near-full-length viral RNA sequences from p24<sup>+</sup> wells. HIV-1 DNA sequences with G to A hypermutations are enclosed in red hashed boxes. The viral outgrowth sequence variants that differ by 1 to 2 nucleotides can be attributed to either ex vivo replication or errors introduced during cDNA synthesis. Average pairwise distances (APDs) calculated by MEGA v6.0 using HIV-1 DNA sequences excluding hypermutated sequences. IS indicates the integration site location of the replicone in the host genome.

The important implications of these findings for clinical management are that changes in the antiretroviral drug regimen or efforts to enhance drug adherence are unlikely to change the viremia that is being produced by infected cell clones, which are not affected by current antiretroviral drugs. Drugs that block virion production from cells that are already infected would be expected to lower viremia, but such agents are not currently available. HIV-1 protease inhibitors that are available block virion maturation into infectious particles, but do not reduce virion production from cells that are already infected (49). Although the frequency of nonsuppressible viremia of clonal origin is unknown, a total of 14 cases referred to us for evaluation, 8 of which are described here, were from a group of approximately 2,000 patients cared for at the referring centers.

In 3 of the 8 donors, identical sequences were found in plasma, proviral DNA, and qVOA, indicating that the viruses were replication competent and potentially transmissible by exposure to blood (Figures 1–3 and Table 3). In a fourth donor, identical sequences were obtained from proviral DNA and qVOA that did not match any viral sequences in plasma, indicating that the viruses in that donor’s plasma had a different origin (Figure 4 and Table 3). In the

remaining 4 donors, there were viruses with identical sequences found in the plasma, but these sequences did not match any of the sequences obtained by qVOA (Supplemental Figures 2–5). In 2 of these 4 donors (K-01 and P-08), proviruses were found that matched the HIV-1 RNA in plasma, but did not match viral RNA sequences from p24<sup>+</sup> qVOA wells (Supplemental Figures 2 and 3 and Table 3). Interestingly, all HIV-1 RNA sequences derived from plasma for donor P-08 were identical, suggesting that nonsuppressible viremia was originating from a single clone. In the other 2 donors, proviruses were not found with sequences that matched HIV-1 RNA in plasma (T-05, 0 of 114 sequences; A-06, 0 of 77 sequences), suggesting that the virus-producing cells were either present at a low frequency in the periphery or absent from the blood and likely residing in lymphoid tissues (Supplemental Figures 3 and 4). Sampling of cells only from the periphery limits the detection of many infected cell clones. In addition, the lack of matches between the virus in the plasma and the viruses recovered from qVOA could be explained by failure to induce the relevant provirus in a single qVOA or that the plasma virus is defective and noninfectious (9, 50).



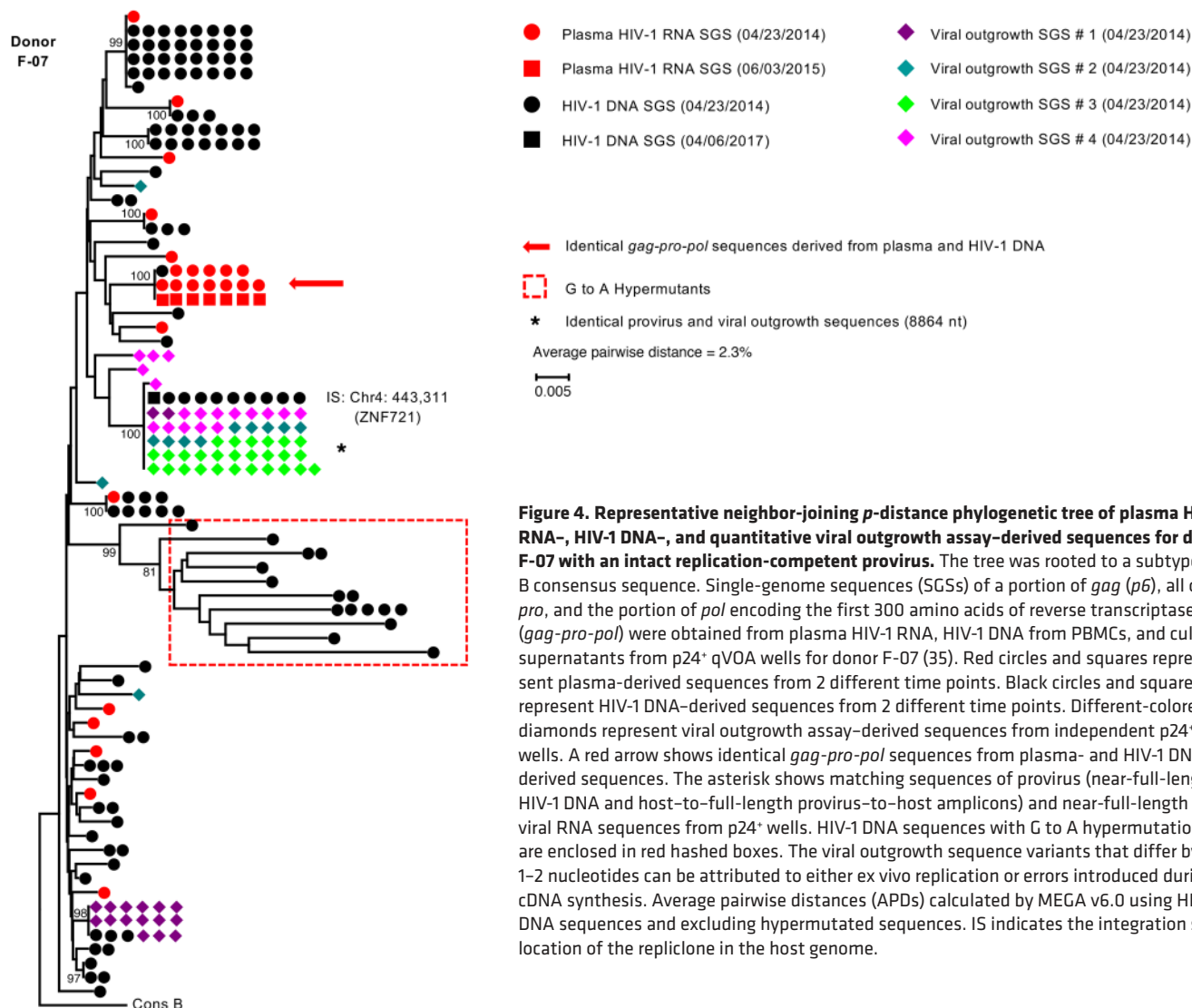
**Figure 3. Representative neighbor-joining *p*-distance phylogenetic tree of plasma HIV-1 RNA-, HIV-1 DNA-, and quantitative viral outgrowth assay-derived sequences from donor C-02 with intact replication-competent proviruses producing nonsuppressible viremia.** The tree was rooted to a subtype B consensus sequence. Single-genome sequences (SGSs) of a portion of *gag* (*p6*), all of *pro*, and the portion of *pol* that encodes the first 300 amino acids of reverse transcriptase (*gag-pro-pol*) were obtained from plasma HIV-1 RNA, HIV-1 DNA from PBMCs, and culture supernatants from p24<sup>+</sup> qVOA wells for donor C-02 (35). Red circles and squares represent plasma-derived sequences from 2 different time points. Black circles and squares represent HIV-1 DNA-derived sequences from 2 different time points. Different-colored diamonds represent viral outgrowth assay-derived sequences from independent p24<sup>+</sup> wells. A red arrow shows identical *gag-pro-pol* sequences for plasma-, HIV-1 DNA-, and viral outgrowth assay HIV-1 RNA-derived sequences. The asterisk shows matching sequences of provirus (near-full-length HIV-1 DNA and host-to-full-length provirus-to-host amplicons) and near-full-length viral RNA sequences from p24<sup>+</sup> wells. HIV-1 DNA sequences with G to A hypermutations are enclosed in red hashed boxes. The viral outgrowth sequence variants that differ by 1 to 2 nucleotides can be attributed to either *ex vivo* replication or errors introduced during cDNA synthesis. Average pairwise distances (APDs) calculated by MEGA v6.0 using HIV-1 DNA sequences and excluding hypermutated sequences. IS indicates the integration site location of the replicone in the host genome.

In the 4 donors with clones containing intact proviruses that produced infectious virus, we identified their corresponding integration sites (Tables 3 and 4 and Supplemental Figure 6). In all cases, the integration sites were within introns of known genes and the proviruses were orientated opposite to the host gene, a pattern that is consistent with what is known about HIV-1 integration sites both in cultured cells and in donors (11, 34, 51). Two of the 4 proviruses were integrated in *ABCA11P* and *ZNF721* within about 17,500 nucleotides of each other, near the end of chromosome 4. Using a larger collection of integration sites obtained from multiple donors on ART, we found no evidence of positive *in vivo* selection for proviruses in either of these 2 overlapping genes (*ABCA11P/ZNF721*) in which the 2 intact proviruses are integrated (data not shown). Thus, their integration in the same overlapping gene is likely to be a coincidence.

ISAs showed that the clones carrying intact proviruses were a small fraction (0.03%–1.1%) of all of the infected cells. However, given the large number of CD4<sup>+</sup> T cells in the human immune system, the clones were estimated to comprise large numbers of infected cells (50 million to 350 million cells)

based on the percentage of matching proviruses and estimated total body CD4<sup>+</sup> T cells (Table 4 and ref. 46). Furthermore and despite the large sizes of these clones, previous work has shown that not all proviruses within a clone are equally expressed, suggesting important differences between daughter cells and the anatomical location, differentiation state, and the microenvironment of the cell that may play an important role in proviral expression (13, 14). The very low frequency of cells infected with the specific intact proviruses makes analyses of the CD4<sup>+</sup> T cell subsets and proviruses involved very challenging given the limitations of cell sampling and current technologies to detect specific proviruses, although such work is in progress.

The reason nonsuppressible viremia developed after years of ART-mediated suppression is not known. The mechanisms that underlie infected-cell expansion and proviral expression are undefined. It has been shown that a cell carrying an intact provirus can have a growth or survival advantage over other cells as a consequence of the location of the integration site (34), and that this advantage may take several years for the clone to become large enough to be clinically manifest. Alternatively, the cell carrying an



**Figure 4. Representative neighbor-joining *p*-distance phylogenetic tree of plasma HIV-1 RNA-, HIV-1 DNA-, and quantitative viral outgrowth assay-derived sequences for donor F-07 with an intact replication-competent provirus.** The tree was rooted to a subtype B consensus sequence. Single-genome sequences (SGSs) of a portion of *gag* (*gag*), all of *pro*, and the portion of *pol* encoding the first 300 amino acids of reverse transcriptase (*gag-pro-pol*) were obtained from plasma HIV-1 RNA, HIV-1 DNA from PBMCs, and culture supernatants from p24<sup>+</sup> qVOA wells for donor F-07 (35). Red circles and squares represent plasma-derived sequences from 2 different time points. Black circles and squares represent HIV-1 DNA-derived sequences from 2 different time points. Different-colored diamonds represent viral outgrowth assay-derived sequences from independent p24<sup>+</sup> wells. A red arrow shows identical *gag-pro-pol* sequences from plasma- and HIV-1 DNA-derived sequences. The asterisk shows matching sequences of provirus (near-full-length HIV-1 DNA and host-to-full-length provirus-to-host amplicons) and near-full-length viral RNA sequences from p24<sup>+</sup> wells. HIV-1 DNA sequences with G to A hypermutations are enclosed in red hashed boxes. The viral outgrowth sequence variants that differ by 1–2 nucleotides can be attributed to either ex vivo replication or errors introduced during cDNA synthesis. Average pairwise distances (APDs) calculated by MEGA v6.0 using HIV-1 DNA sequences and excluding hypermutated sequences. IS indicates the integration site location of the replicone in the host genome.

intact provirus encounters its cognate antigen, becomes activated and proliferates, and its provirus becomes expressed. To produce sustained viremia the cognate antigen would likely have to persist, suggesting its source may be a chronic infection (e.g., cytomegalovirus, Epstein-Barr virus, or HIV-1 itself) or host derived, as has been observed in one instance in a person with malignancy (10).

Our findings have important implications for efforts to develop a cure for HIV-1 infection. It is widely believed that the latent HIV-1 reservoir is the main obstacle to a cure. Some have questioned the relevance of persistent viremia on ART, but our findings show that some of the viruses that persist in plasma can be infectious and could quickly initiate viral rebound if ART is discontinued (48, 52, 53). This “active” HIV-1 reservoir is thus also a key barrier to curing HIV-1 infection. It is possible, in cases in which there are much lower levels of viremia, that there are smaller clones of infected cells releasing infectious virus at levels that are below detection in plasma, but are still capable of rekindling viral replication if ART is interrupted. Previous studies using analytical treatment interruption (ATI) have observed a discordance between the rebounding viruses and the replication-competent virus that was induced

in a VOA (54, 55). This discordance may be due to differences in the mechanisms of proviral activation in vivo versus ex vivo (i.e., qVOA conditions), inherent sampling limitations of VOAs, and recombination between rebounding variants in vivo that generates new variants. Some of the donors in the current study are already viremic with infectious clonal virus that could rapidly seed new rounds of replication following the interruption of treatment, although ATI was not performed in the current study.

The nonsuppressible viremia in our study participants could result in immune-activating or inflammatory effects, although recent work from Gandhi et al. has shown no association between measures of HIV-1 persistence and inflammation/activation in over 300 well-suppressed individuals (33). Additional studies with matched controls without nonsuppressible viremia are needed to address the role of immune activation and inflammation in clonal expansion and virus production from clones.

Importantly, the mechanisms that allow large clones to expand and continually produce virus need to be defined. At present, we cannot determine whether there is a small fraction of each replicone that produces virus continuously, somehow evad-

**Table 3. Single-genome sequence matches between plasma HIV-1 RNA, proviruses, and viral outgrowth cultures**

Donor ID	Identical Plasma HIV-1 RNA Sequences <sup>A</sup>	Proviral Sequences Matching Dominant Plasma HIV-1 RNA Sequences (% matching) <sup>B</sup>	qVOA p24 <sup>+</sup> Wells with Matching Sequences in Plasma HIV-1 RNA <sup>C,D</sup>	Proviral Sequences Matching Dominant qVOA HIV-1 RNA Sequences (% matching) <sup>E</sup>
R-09	50.0%	10.7%	100% (10/10)	10.7%
C-03	63.3%	2.9%	66.7% (2/3)	2.9%
C-02	90.7%	7.5%	100% (3/3)	7.5%
F-07	53.8%	0.3%	0% (0/4)	3.0%
K-01	37.5%	3.9%	0% (0/4)	<3.9%
P-08	100%	8.0%	0% (0/3)	<8.0%
T-05	76.7%	<0.7%	0% (0/3)	<0.7%
A-06	48.4%	<0.7%	0% (0/2)	<0.7%

<sup>A</sup>Percentage of all HIV-1 RNA single-genome sequences from plasma (1.3 kb of *gag-pro-pol*) that are identical or differ by 1 to 2 bases from the group of identical sequences. <sup>B</sup>Identical proviruses (single-genome sequences) in PBMCs that match HIV-1 RNA sequences from plasma (1.3 kb of *gag-pro-pol*). <sup>C</sup>Quantitative viral outgrowth (qVOA) HIV-1 RNA single-genome sequences (1.3 kb of *gag-pro-pol*). <sup>D</sup>Percentage and number of *gag* p24 antigen-positive (p24<sup>+</sup>) qVOA wells with identical sequences that match plasma HIV-1 RNA sequences (1.3 kb of *gag-pro-pol*). <sup>E</sup>Percentage of total proviral HIV-1 DNA single-genome sequences that match identical qVOA HIV-1 RNA single-genome sequences (1.3 kb of *gag-pro-pol*). Bold indicates that the proviruses in clonally expanded cells are intact (replication-competent) and either contribute to persistent infectious viremia (C-02, C-03, R-09) or can be induced ex vivo (F-07) to produce infectious virus (qVOA).

ing killing by the virus or the immune response, or whether the virus-producing cells die after their proviruses are expressed only to be replaced by equal numbers of cells that are able to produce virus. Similarly, the fraction of each replicone that is producing virions needs to be defined, but methodological approaches to investigate this question require development. It is also uncertain if these clones can be recognized by either humoral or cellular immunological responses. Detailed analyses of viral susceptibility to autologous plasma and a panel of monoclonal antibodies is in progress, as is assessment of mutations that could result in CTL escape. The large size of the clones we found that carry replication-competent proviruses suggest that developing therapies to eliminate them or control their ability to produce virus represents a formidable challenge.

## Methods

**Study design.** The primary objective of the study was to evaluate non-suppressible viremia and confirm its clonal cellular origin. The schematic in Supplemental Figure 1 provides an overview of the multistep methodological approach to evaluating non-suppressible viremia and

confirming its clonal cellular origin. This included SGS of amplicons containing HIV-1 *gag-pro-pol* and/or NFL genomes derived from (a) plasma, (b) PBMCs or total CD4<sup>+</sup> T cells, and/or (c) p24<sup>+</sup> qVOA wells in 8 of 8 participants. Furthermore, ISAs and host-to-full-length provirus-to-host amplification and sequencing were used to confirm clonality in 4 of the 8 participants. A more detailed description of methods used is provided below.

**Participants and sample collection.** Study participants were referred from the University of Pittsburgh Medical Center HIV-AIDS Program or from the Allegheny Health Network Positive Health Clinic and enrolled into the study at the University of Pittsburgh Clinical Trials Unit between April of 2014 and June of 2015. Inclusion criteria for the study were (a) initial suppression of plasma HIV-1 RNA to below the limit of detection of commercial assays (<20 or <40 copies/mL) on a DHHS-recommended ART regimen; (b) followed by clinically detectable plasma viremia (HIV-1 RNA >20 copies/mL) for more than 6 months, as measured by the COBAS Ampliprep/COBAS TaqMan, v2.0 assay (TMv2.0) (Roche) or the m2000sp/RealTime HIV-1 Assay (Abbott); and (c) assessment by the referring physician that the patient was fully compliant with their ART regimen. Switches in ART regimen

**Table 4. Integration site analyses of intact replication-competent proviruses in expanded clones**

Donor	Intact Provirus Integration Site <sup>A</sup>	Gene <sup>B</sup>	Location <sup>C</sup>	Proviral Orientation Relative to Gene <sup>D</sup>	Frequency of Clone by Integration Site Assay <sup>E</sup>	Estimated Size of Clone (Number of Cells) <sup>F</sup>
R-09	Chr4: 425,766	ABCA11P	Intron	-	0.03% (3/9939)	3.5 × 10 <sup>8</sup>
C-03	Chr12: 133,777,660	ZNF268	Intron	-	1.10% (44/4164)	1.3 × 10 <sup>8</sup>
C-02	Chr5: 138,660,320	MATR3	Intron	-	0.65% (3/459)	5.2 × 10 <sup>7</sup>
F-07	Chr4: 443,311	ZNF721/ABCA11P	Intron	-	0.05% (4/8407)	8.0 × 10 <sup>7</sup>

<sup>A</sup>Chromosomal location of the proviral 3' LTR in the human genome hg19 reference using UCSC Genome Browser (44, 45). <sup>B</sup>Gene containing the provirus. <sup>C</sup>Intragenic location of provirus. <sup>D</sup>Proviral orientation relative to gene: positive (+) in same orientation as gene; negative (-) in opposite orientation to gene. <sup>E</sup>Frequency of integrant relative to total integrants determined by integration site analyses (34). <sup>F</sup>Estimated size of each clone derived from the frequency of matching proviral sequences in CD4<sup>+</sup> T cells frequency and the total number of CD4<sup>+</sup> T cells in each donor. Total CD4<sup>+</sup> cells calculated from percentage of CD4<sup>+</sup> cells × total body lymphocytes (2 × 10<sup>12</sup>) (46).



or intensification with another antiretroviral did not exclude participants from the study. Follow-up of study participants after evaluation for nonsuppressible viremia included continuation of primary care with immunologic and virologic monitoring.

Longitudinal samples were collected at 2 or more time points as large-volume phlebotomy (100–180 mL) or plasmapheresis and leukapheresis between April of 2014 and April of 2019. Plasma, PBMCs, and total CD4<sup>+</sup> T cells were isolated and stored as previously reported (22). Data from these samples were collected between May of 2014 and July of 2019 and included 2 or more time points for each participant.

**Quantification of plasma HIV-1 RNA and cellular HIV-1 DNA and RNA.** HIV-1 DNA and RNA in PBMCs or total CD4<sup>+</sup> T cells was quantified by qPCR targeting the 3' end of integrase, as previously reported (22).

**qVOA.** qVOAs were performed using total CD4<sup>+</sup> T cells. The frequency of HIV-1-infected cells carrying an IUPM was determined by a maximum likelihood method, as reported previously (8, 23, 24).

**SGS.** SGS of *gag* (*p6*), *pro*, and the portion of *pol* encoding the first 300 amino acids of reverse transcriptase (*gag-pro-pol*) was performed using HIV-1 RNA from plasma or from p24<sup>+</sup> qVOA wells or from cellular HIV-1 DNA as follows: (a) endpoint dilution of extracted nucleic acid to single HIV-1 template per PCR reaction as determined by Poisson distribution statistics, (b) generation of an approximately 1.56 kb RT-PCR or PCR amplicon, and (c) bidirectional sequencing of the amplicon by the Sanger method (8, 35). The following modifications to SGS were performed: donors with plasma HIV-1 RNA levels less than 200 copies/mL, all of the extracted RNA was used for cDNA synthesis, cDNA was diluted 1:2 or 1:4 in 5 mM Tris-HCl, and all diluted cDNA was used to seed multiple PCR reactions.

**NFL single proviral and viral genome amplification and sequencing.** NFL HIV-1 DNA was amplified by nested PCR from genomic DNA extracted from PBMCs or total CD4<sup>+</sup> T cells at a proviral endpoint of a single template per PCR reaction as determined by Poisson distribution statistics. PCR amplifications were performed using the 2× RANGER DNA Polymerase Mix according to the manufacturer's recommendations (Bioline) and previously reported primers (56). The sizes of the NFL amplicons were confirmed with the PerkinElmer GX Touch 24 LabChip bioanalyzer using the 12K DNA module. NFL amplicons containing 16 bp symmetrical barcodes were size selected by BluePippin (Sage Science) and libraries were constructed using the PacBio SMRTbell Template Prep Kit before PacBio sequencing (Pacific Biosciences). Alternatively, NFL amplicons were sequenced by Illumina MiSeq using the KAPA HyperPlus kit (KAPA Biosystems) for library construction and the MiSeq nano v2 500 cycle, 2 × 250 run kit with dual index (Illumina) for sequencing. SGS of overlapping half genomes from virion-associated HIV-1 RNA in p24<sup>+</sup> qVOA wells was performed as reported previously (8, 36).

**Assessment of cell clonality.** A clone of HIV-1-infected cells was defined by identifying multiple cells that had identical proviral sequences integrated into the identical position in the human genome. The methods used to confirm clonality are described below.

**MDA of genomic DNA.** Whole cellular genome amplification was done as previously described (43) with the following modifications: (a) genomic DNA was diluted to a proviral endpoint of 1 template per reaction; (b) reactions volumes were 25 μL and contained denaturing solution (final concentrations: 24.4 mM KOH and 20 μM random hexamers [OH-5'-NNNN\*N-N-3']), neutralization solution (final concentrations: 25 mM Trizma-HCl, pH 7.5 and 600 mM trehalose), and

reaction mix (final concentrations: 1× Phi buffer, 20 mM KCl, 2 mM DTT, 100 mM BSA, 1.8 mM dNTPs, 600 mM trehalose, and 1 unit phi29 DNA polymerase); and (c) incubations were performed at 40°C for 20 hours and terminated at 65°C for 10 minutes.

**ISAs.** ISAs were conducted as previously described (34). For proviruses with specific viral sequences of interest (i.e., matching plasma HIV-1 RNA sequences), ISA was performed with the following modifications: the starting template was 0.8× SPRI-purified MDA material that contained the *gag-pro-pol* sequence that matched the plasma HIV-1 RNA or p24<sup>+</sup> qVOA HIV-1 RNA sequence of interest. ISA was performed using an in-house workflow using MDA and a specificity-enhancing linker-mediated PCR that amplifies across the 5'LTR host/virus junction (37).

**Host-to-full-length provirus-to-host amplification and sequencing.** For specific clones of interest, full-length sequences (host-to-full-length provirus-to-host) of the proviruses integrated at the same site were assessed for identity. Proviral integration site positions were based on data in the Human Genome Browser at UCSC using the human genome assembly reference Hg19 sequence (44, 45). These sequences were also used to design primers for the human host sequences adjacent to full-length proviruses. Host-specific primers were used with SGS *gag-pro-pol* proviral specific primers to amplify the entire provirus and the flanking 5' and 3' host sequences as 2 overlapping fragments from genomic DNA extracted from PBMCs or total CD4<sup>+</sup> T cells using primers specific for each integrant of interest (Supplemental Table 4). The amplification was performed using 2× RANGER DNA Polymerase Mix as described above for NFL proviral amplification. The products were sequenced using either Sanger or the Illumina platform. Primers used for Sanger sequencing are listed in Supplemental Table 5.

**Sequence alignments, quality control, and phylogenetic analyses.** Sequence alignments and phylogenetic analyses of the *gag-pro-pol* sequences were performed as reported previously (8). This included alignments, exclusion of mixtures, and quality control of sequences in Sequencher v5.0 (Gene Codes) (Supplemental Figure 7). Hypermutant sequences were determined using the Stanford University HIV Drug Resistance Database HIVdb program (<https://hivdb.stanford.edu/hivdb/by-sequences>) (38). Neighbor-joining *p*-distance phylogenetic trees were rooted to subtype B with bootstrapping at 1,000 replicates per tree using MEGA6 (57). MEGA6 was also used to calculate the APD by including only nonhypermutated HIV-1 DNA sequences in the analyses.

**HIV-1 subtype, drug susceptibility analyses, drug concentration determination, and coreceptor tropism.** HIV-1 subtype- and genotype-predicted susceptibilities to antiretroviral drugs were determined by the HIVdb program (Stanford University HIV Drug Resistance Database) (38). Drug level concentrations in human plasma for the non-nucleoside reverse transcriptase inhibitor efavirenz and the protease inhibitors darunavir, atazanavir, and ritonavir were measured by a gradient separation through ultra performance liquid chromatography with electrospray ionization tandem mass spectrometry for detection (39). Drug concentration in human plasma for the integrase inhibitor dolutegravir was measured using a protein precipitation method with isocratic separation by liquid chromatography with tandem mass spectrometry for detection (40). Coreceptor tropism was determined by the Geno2Pheno bioinformatics software using both established and individually selected cutoffs (41).

**Immunophenotyping.** Surface flow cytometric staining of CD3, CD4, CD8, CD19, CD56, CD25, CD69, CD38, HLA-DR, and CD107a

(BD Biosciences) was performed on a BD LSRII cytometer according to standard published methods (58), analyzed using FlowJo (BD), and results compared with published results of healthy adults (25–32).

**Calculating the size of the repliclones.** The calculations used to estimate the size of each repliclone were based on (a) the estimated total number of lymphocytes found in the human body, a value of  $2 \times 10^{12}$  cells (46); (b) the HIV-1 DNA copies/million CD4<sup>+</sup> T cells (22); and (c) the fraction of proviruses that matched the relevant repliclone, as determined by SGS. Repliclone copies/ $10^6$  CD4<sup>+</sup> cells was determined by multiplying HIV-1 DNA/ $10^6$  CD4<sup>+</sup> cells by the frequency of the repliclone. The number of CD4<sup>+</sup> cells per donor was determined by multiplying the percentage of CD4<sup>+</sup> T cells by  $2 \times 10^{12}$  lymphocytes in the human body (46). The size each repliclone was then determined by multiplying the number of repliclone copies/million CD4<sup>+</sup> T cells by the estimated total body CD4<sup>+</sup> T cells. This calculation does not account for skewed distribution of CD4<sup>+</sup> T cells in nonlymphoid tissues.

**Data availability.** Sequences have been submitted to the GenBank database (accession numbers MT743286–MT744356 and MT744357–MT745577). Viral sequences and integration sites were also submitted to the Proviral Sequence Database (PSD) and the Retrovirus Integration Database (RID; <https://rid.ncifcrf.gov/>) at the US Department of Health and Human Services, NIH, National Cancer Institute (59). Next-generation sequencing data from PacBio and Illumina platforms were submitted to the NIH Sequence Read Archive (submission ID: SUB7769720; BioProjectID: PRJNA646243; accession numbers: SAMN15534104–SAMN15534127). Sequences were also submitted to the database at the National Cancer Institute, HIV Dynamics and Replication Program (<https://fscigl-trups01.ncifcrf.gov/index.php>) referenced by the identifier DRP ID 4.

**Statistics.** Average pairwise distances were calculated by MEGA6 through the exclusion of hypermutants and identification of hypermutants was performed using the HIVdb algorithm (Stanford University HIV Drug Resistance Database) (38, 57).

**Study approval.** The University of Pittsburgh Institutional Review Board approved the study. The clinical research staff completed enrollment procedures and all participants provided written informed consent before inclusion in the study.

## Author contributions

JWM conceived and designed the research and was responsible for oversight of human subject research. EKH, KWJ, LDB, MFK, JMC, JWR, XW, SHH, and JWM provided intellectual input. MDS was responsible for processing of plasma and cells from donors. EKH,

KWJ, LDB, SG, MDS, JIJ, CT, JKB, JCC, BFK, GDM, MJB, and WS performed experiments and/or developed new reagents/tools. EKH, KWJ, LDB, WS, MFK, JMC, JWR, XW, SHH, and JWM analyzed data. EKH, SHH, and JWM wrote and critically revised the manuscript.

## Acknowledgments

We thank the individuals who volunteered for this study. We also thank the referring physicians, Bernard Macatangay and Sharon Riddler, from the University of Pittsburgh Medical Center (UPMC), and Chiu-Bin Hsiao from Allegheny Health Network. We are grateful for the efforts of the clinical staff at the UPMC HIV-AIDS Program, especially Renee Weinman and Jamie Ideluca. We are also thankful for the efforts of Lorraine Pollini for proofreading, formatting, and submission of this manuscript. This project has been supported in part by funding to JWM, BFK, MFK, JWR, XW, and SHH by the Intramural Research Programs of the National Cancer Institute, NIH, under contract numbers HHSN261200800001E and 75N91019D00024. JKB receives grant support from the Howard Hughes Medical Research Fellows Program, Howard Hughes Medical Institute. JWM also receives grant support from the Bill & Melinda Gates Foundation award OPP1115715 and the National Institute for Allergy and Infectious Diseases, NIH, to the I4C Martin Delaney Collaboratory contract number UM1AI126603. MFK and JWR receive grant support from the Office of AIDS Research. JMC receives grant support from The American Cancer Society and the National Cancer Institute through a Leidos subcontract number I3XS110. GDM receives grant support from the University of Rochester Center for AIDS Research, contract number P30AI078498, and from the University of Rochester HIV/AIDS Clinical Trials Unit, contract number UM1 AI069511-08. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views or policies of the National Institute of Allergy and Infectious Diseases, the NIH, or the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

Address correspondence to: John W. Mellors, Division of Infectious Diseases, Department of Medicine, University of Pittsburgh, 3550 Terrace Street, Scaife Hall, Suite 818, Pittsburgh, Pennsylvania 15261, USA. Email: [jwm1@pitt.edu](mailto:jwm1@pitt.edu).

1. Palella FJ, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med.* 1998;338(13):853–860.
2. Finzi D, et al. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science.* 1997;278(14):1295–1300.
3. Wong JK, et al. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science.* 1997;278(14):1291–1295.
4. Siliciano JD, et al. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4<sup>+</sup> T cells. *Nat Med.* 2003;9(6):727–728.
5. Ho YC, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell.* 2013;155(3):540–551.
6. Bruner K, et al. Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat Med.* 2016;22(9):1043–1049.
7. Lorenzi JC, et al. Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA. *Proc Natl Acad Sci U S A.* 2016;113(49):E7908–E7916.
8. Bui JK, et al. Ex vivo activation of CD4<sup>+</sup> T-cells from donors on suppressive ART can lead to sustained production of infectious HIV-1 from a subset of infected cells. *PLoS Pathog.* 2017;13(2):e1006230.
9. Hosmane NN, et al. Proliferation of latently infected CD4<sup>+</sup> T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J Exp Med.* 2017;214(4):959–972.
10. Simonetti FR, et al. Clonally expanded CD4<sup>+</sup> T cells can produce infectious HIV-1 in vivo. *Proc Natl Acad Sci U S A.* 2016;113(7):1883–1888.
11. Einkauf KB, et al. Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy. *J Clin Invest.* 2019;129(3):988–998.
12. Kearney MF, et al. Origin of rebound plasma HIV includes cells with identical proviruses that are transcriptionally active before stopping of antiret-

- roviral therapy. *J Virol*. 2015;90(3):1369–1376.
13. Wiegand A, et al. Single-cell analysis of HIV-1 transcriptional activity reveals expression of proviruses in expanded clones during ART. *Proc Natl Acad Sci U S A*. 2017;114(18):E3659–E3668.
  14. Musick A, et al. HIV infected T cells can proliferate in vivo without inducing expression of the integrated provirus. *Front Microbiol*. 2019;10:2204.
  15. McManus WR, et al. HIV-1 in lymph nodes is maintained by cellular proliferation during antiretroviral therapy. *J Clin Invest*. 2019;129(11):4629–4642.
  16. Palmer S, et al. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol*. 2003;41(10):4531–4536.
  17. Cillo AR, et al. Improved single-copy assays for quantification of persistent HIV-1 viremia in patients on suppressive antiretroviral therapy. *J Clin Microbiol*. 2014;52(11):3944–3951.
  18. Tosiano MA, Jacobs JL, Shutt KA, Cyktor JC, Mellors JW. A simpler and more sensitive single-copy HIV-1 RNA assay for quantification of persistent HIV-1 viremia in individuals on suppressive antiretroviral therapy. *J Clin Microbiol*. 2019;57(3):e01714-18.
  19. Dinoso JB, et al. Treatment intensification does not reduce residual HIV-1 viremia in patients on highly active antiretroviral therapy. *Proc Natl Acad Sci U S A*. 2009;106(23):9403–9408.
  20. McMahon D, et al. Short-course raltegravir intensification does not reduce persistent low-level viremia in patients with HIV-1 suppression during receipt of combination antiretroviral therapy. *Clin Infect Dis*. 2010;50(6):912–919.
  21. Gandhi RT, et al. The effect of raltegravir intensification on low-level residual viremia in HIV-infected patients on antiretroviral therapy: a randomized controlled trial. *PLoS Med*. 2010;7(8):e1000321.
  22. Hong F, et al. Novel assays for measurement of total cell-associated HIV-1 DNA and RNA. *J Clin Microbiol*. 2016;54(4):902–991.
  23. Siliciano JD, Siliciano RF. Enhanced culture assay for detection and quantitation of latently infected, resting CD4<sup>+</sup> T-cells carrying replication-competent virus in HIV-1-infected individuals. *Methods Mol Biol*. 2005;304:3–15.
  24. Rosenbloom DI, Elliott O, Hill AL, Henrich TJ, Siliciano JM, Siliciano RF. Designing and interpreting limiting dilution assays: general principles and applications to the latent reservoir for human immunodeficiency virus-1. *Open Forum Infect Dis*. 2015;2(4):ofv123.
  25. Bisset LR, Lung TL, Kaelin M, Ludwig E, Dubs RW. Reference values for peripheral blood lymphocyte phenotypes applicable to the healthy adult population in Switzerland. *Eur J Haematol*. 2004;72(3):203–212.
  26. Yi JS, et al. Establishment of normative ranges of the healthy human immune system with comprehensive polychromatic flow cytometry profiling. *PLoS One*. 2019;14(12):e0225512.
  27. Waschbisch A, et al. Analysis of CD4<sup>+</sup> CD8<sup>+</sup> double-positive T cells in blood, cerebrospinal fluid and multiple sclerosis lesions. *Clin Exp Immunol*. 2014;177(2):404–411.
  28. Overgaard NH, Jung J-W, Steptoe RJ, Wells JW. CD4<sup>+</sup>/CD8<sup>+</sup> double-positive T cells: more than just a developmental stage? *J Leukoc Biol*. 2015;97(1):31–38.
  29. Ahmad F. High frequencies of polyfunctional CD8<sup>+</sup> NK cells in chronic HIV-1 infection are associated with slower disease progression. *J Virol*. 2014;88(21):12397–12408.
  30. Michou VI, Kanavaros P, Athanassiou V, Chronis GB, Stabamas S, Tsilivakos V. Fraction of the peripheral blood concentration of CD56<sup>+</sup>/CD16<sup>+</sup>/CD3<sup>+</sup> cells in total natural killer cells as an indication of fertility and infertility. *Fertil Steril*. 2003;80 Suppl 2:691–697.
  31. Kilmartin DJ, Fletcher ZJ, Almeida JA, Liversidge J, Forrester JV, Dick AD. CD69 expression on peripheral CD4<sup>+</sup> T cells parallels disease activity and is reduced by mycophenolate mofetil therapy in uveitis. *Invest Ophthalmol Vis Sci*. 2001;42(6):1285–1292.
  32. Giacoia-Gripp CBW, et al. Changes in the NK cell repertoire related to initiation of TB treatment and onset of immune reconstitution inflammatory syndrome in TB/HIV co-infected patients in Rio de Janeiro, Brazil-ANRS 12274. *Front Immunol*. 2019;10:1800.
  33. Gandhi RT, et al. Levels of HIV-1 persistence on antiretroviral therapy are not associated with markers of inflammation or activation. *PLoS Pathog*. 2017;13(4):e1006285.
  34. Maldarelli F, et al. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science*. 2014;345(6193):179–183.
  35. Palmer S, et al. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J Clin Microbiol*. 2005;43(1):406–413.
  36. Salazar-Gonzalez JF, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J Exp Med*. 2009;206(6):1273–1289.
  37. Joseph KW, et al. High-throughput sequencing of integrated HIV-1 reveals novel proviral structures. Abstract presented at: 9th International Workshop on HIV Persistence during Therapy; December 10–13, 2019; Miami, Florida, USA. <https://www.hiv-persistence.com/wp-content/uploads/2019/12/Programme-final-A5-WORKSHOP-Miami-2019-USA-10-12-1.pdf>. Accessed August 24, 2020.
  38. Liu TF, Shafer RW. Web resources for HIV type 1 genetic-resistance test interpretation. *Clin Infect Dis*. 2006;42(11):1608–1618.
  39. Kiser JJ, et al. Boceprevir and Antiretroviral Pharmacokinetic Interactions in HIV/HCV Co-infected Persons: AIDS Clinical Trials Group Study A5309s. *Drugs R D*. 2017;17(4):557–567.
  40. Bennetto-Hood C, Tabolt G, Savina P, Acosta EP. A sensitive HPLC-MS/MS method for the determination of dolutegravir in human plasma. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2014;945–946:225–232.
  41. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol*. 2007;25(12):1407–1410.
  42. Wagner TA, et al. HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science*. 2014;345(6196):570–573.
  43. Pan X, et al. A procedure for highly specific, sensitive, and unbiased whole-genome amplification. *Proc Natl Acad Sci U S A*. 2008;105(40):15499–15504.
  44. Kent WJ, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996–1006.
  45. Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res*. 2002;12(4):656–664.
  46. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Molecular Biology of the Cell*, 4th ed. Garland Science; 2002.
  47. Tobin NH, et al. Evidence that low-level viremia during effective highly active antiretroviral therapy result from two processes: expression of archival virus and replication of virus. *J Virol*. 2005;79(15):9625–9634.
  48. Bailey JR, et al. Residual human immunodeficiency virus type 1 viremia in some patients on antiretroviral therapy is dominated by a small number of invariant clones rarely found in circulating CD4<sup>+</sup> T cells. *J Virol*. 2006;80(13):6441–6645.
  49. Adamson CS. Protease-mediated maturation of HIV: inhibitors of protease and the maturation process. *Mol Biol Int*. 2012;604261:1–13.
  50. Rassler S, Ramirez R, Khoury N, Skowron G, Sahu GK. Prolonged persistence of a novel replication-defective HIV-1 variant in plasma of a patient on suppressive therapy. *Virol J*. 2016;13(157):1–13.
  51. Schroder ARW, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*. 2002;110(4):521–529.
  52. Joos B, et al. HIV rebounds from latently infected cells, rather than from continuing low-level replication. *Proc Natl Acad Sci U S A*. 2008;105(3):16725–16730.
  53. De Scheerder MA, et al. HIV rebound is predominantly fueled by genetically identical viral expansions from diverse reservoirs. *Cell Host Microbe*. 2019;26(3):347–358.e7.
  54. Cohen YZ, et al. Relationship between latent and rebound viruses in a clinical trial of anti-HIV-1 antibody 3BNC117. *J Exp Med*. 2018;215(9):2311–2324. 2324.
  55. Mendoza P, et al. Combination therapy with anti-HIV-1 antibodies maintains viral suppression. *Nature*. 2018;561(7724):479–483.
  56. Li B, et al. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. *J Virol*. 2006;81(1):193–201.
  57. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 2013;30(12):2725–2729.
  58. Garcia-Bates TM, et al. Contrasting roles of the PD-1 signaling pathway in dendritic cell-mediated induction and regulation of HIV-1-specific effector T cell functions. *J Virol*. 2019;93(5):e02035-18.
  59. Shao W, et al. HIV Proviral Sequence Database: a new public database for near full-length HIV proviral sequences and their meta-analyses. *AIDS Res Hum Retroviruses*. 2019;00(1):1–2.