

HIV-1 proviral landscapes distinguish posttreatment controllers from noncontrollers

Radwa Sharaf,^{1,2} Guinevere Q. Lee,³ Xiaoming Sun,³ Behzad Etemad,¹ Layla M. Aboukhater,¹ Zixin Hu,¹ Zabrina L. Brumme,^{4,5} Evgenia Aga,⁶ Ronald J. Bosch,⁶ Ying Wen,¹ Golnaz Namazi,¹ Ce Gao,³ Edward P. Acosta,⁷ Rajesh T. Gandhi,^{3,8} Jeffrey M. Jacobson,⁹ Daniel Skiest,¹⁰ David M. Margolis,¹¹ Ronald Mitsuyasu,¹² Paul Volberding,¹³ Elizabeth Connick,¹⁴ Daniel R. Kuritzkes,¹ Michael M. Lederman,¹⁵ Xu G. Yu,^{3,7} Mathias Lichterfeld,^{1,3} and Jonathan Z. Li¹

¹Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. ²Harvard PhD Program in Virology, Division of Medical Sciences, Harvard University, Boston, Massachusetts, USA.

³Ragon Institute of MGH, MIT, and Harvard, Cambridge, Massachusetts, USA. ⁴Simon Fraser University, Burnaby, British Columbia, Canada. ⁵British Columbia Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada. ⁶Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA. ⁷University of Alabama at Birmingham, Birmingham, Alabama, USA. ⁸Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA. ⁹Temple University, Philadelphia, Pennsylvania, USA. ¹⁰University of Massachusetts Medical School-Baystate, Springfield, Massachusetts, USA. ¹¹UNC School of Medicine, Chapel Hill, North Carolina, USA. ¹²UCLA, Los Angeles, California, USA. ¹³Gladstone Center for AIDS Research, UCSF, San Francisco, California, USA. ¹⁴University of Arizona College of Medicine, Tucson, Arizona, USA. ¹⁵Case Western Reserve University, Cleveland, Ohio, USA.

HIV posttreatment controllers (PTCs) represent a natural model of sustained HIV remission, but they are rare and little is known about their viral reservoir. We obtained 1,450 proviral sequences after near-full-length amplification for 10 PTCs and 16 posttreatment noncontrollers (NCs). Before treatment interruption, the median intact and total reservoir size in PTCs was 7-fold lower than in NCs, but the proportion of intact, defective, and total clonally expanded proviral genomes was not significantly different between the 2 groups. Quantification of total but not intact proviral genome copies predicted sustained HIV remission as 81% of NCs, but none of the PTCs had a total proviral genome greater than 4 copies per million peripheral blood mononuclear cells (PBMCs). The results highlight the restricted intact and defective HIV reservoir in PTCs and suggest that total proviral genome burden could act as the first biomarker for identifying PTCs. Total and defective but not intact proviral copy numbers correlated with levels of cell-associated HIV RNA, activated NK cell percentages, and both HIV-specific CD4⁺ and CD8⁺ responses. These results support the concept that defective HIV genomes can lead to viral antigen production and interact with both the innate and adaptive immune systems.

Introduction

Despite years of suppressive antiretroviral therapy (ART), the vast majority of individuals infected with HIV will experience rapid viral rebound during an analytic treatment interruption (ATI) (1). Efforts to achieve an HIV cure were boosted by the case of the Berlin patient who achieved an apparently sterilizing cure with no detectable virus after undergoing hematopoietic stem cell transplantation with donor cells that were naturally resistant to HIV infection (2). However, this approach has limited generalizability given the substantial mortality associated with stem cell transplantation and the rarity of donors with cells resistant to HIV. An alternative approach is a functional cure, or sustained HIV remission, where proviral HIV DNA may still be detectable, but patients maintain viral suppression even after ART discontinuation. These individuals are termed HIV posttreatment controllers (PTCs).

The most comprehensive description of PTCs so far has been the VISCONTI cohort of 14 PTCs (3). These individuals were treated during early HIV infection, but unlike HIV elite controllers (ECs), favorable HLA alleles associated with viral control were not over-represented in these participants. However, PTCs are exceedingly rare and the HIV proviral reservoir determinants of posttreatment control are largely unexplored. We addressed this question using a group of PTCs identified from previously completed AIDS Clinical Trials Group (ACTG) studies (4).

HIV DNA and cell-associated RNA (CA-RNA) were detected before ATI in PTCs from both the VISCONTI (3) and ACTG (4) participants. Currently, a smaller HIV reservoir size is thought to be a contributing factor for HIV posttreatment control (5, 6). However, PCR-based methods of HIV reservoir quantification of proviral DNA overestimate the true reservoir size, as most proviral genomes are defective (7–10). We hypothesized that individuals with a smaller intact proviral reservoir prior to treatment interruption are more likely to be PTCs. To explore this question, we performed near-full-length sequencing of proviral genomes using next-generation sequencing of single-genome amplicons for a group of PTCs and posttreatment noncontrollers (NCs). This approach allows for an in-depth assessment of the intact and defective HIV reservoir, as well as providing evidence of likely clonal expansion. Assessing the types of defective proviruses

Conflict of interest: JZL has received research support and consulted for Gilead and Merck. RTG has received educational grants from Gilead, Merck, and ViiV. DRK serves as a consultant to and/or has received research support from Gilead, Janssen, Merck, and ViiV. DS has received research support from Gilead, ViiV, and GlaxoSmithKline. MML has received research support from Gilead.

Submitted: February 15, 2018; **Accepted:** July 3, 2018.

Reference information: *J Clin Invest.* 2018;128(9):4074–4085.

<https://doi.org/10.1172/JCI120549>.

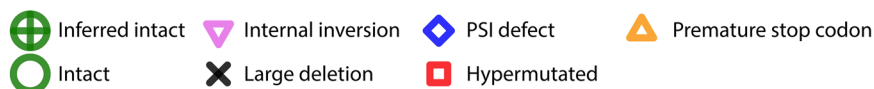
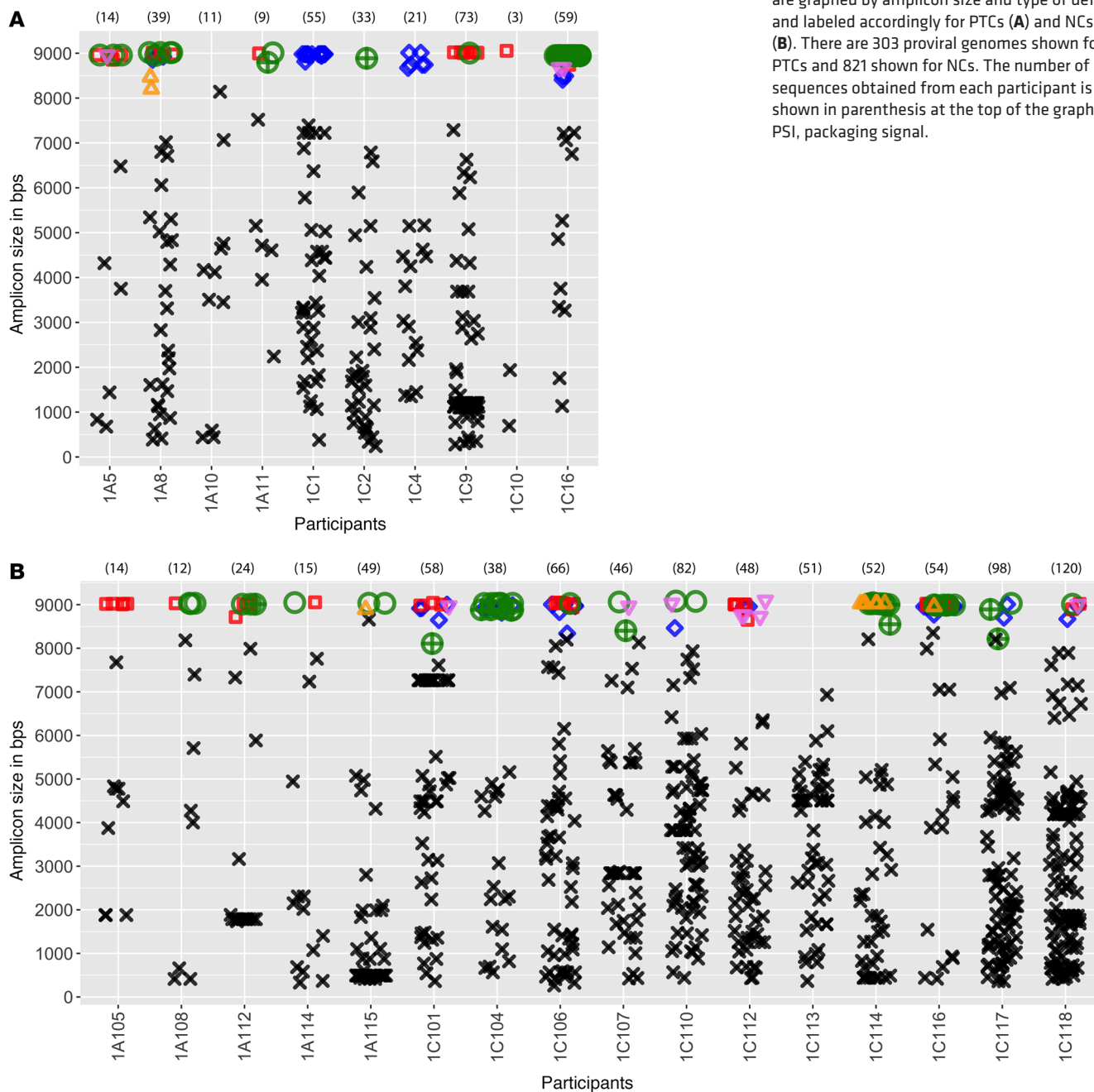


Figure 1. Composition of the proviral landscape pretreatment interruption. (A and B) For each participant, the amplified proviral sequences are graphed by amplicon size and type of defect and labeled accordingly for PTCs (A) and NCs (B). There are 303 proviral genomes shown for PTCs and 821 shown for NCs. The number of sequences obtained from each participant is shown in parenthesis at the top of the graph. PSI, packaging signal.



present (e.g., deleted, hypermutated) provides insight into potential mechanisms underlying posttreatment HIV control. Furthermore, we assessed the relationship among different proviral species, levels of intracellular HIV RNA expression, and host immune responses to define the functional capacity of the proviral reservoir and immune correlates of reservoir size.

Results

Overview of study population and proviral sequencing. Ten PTCs were identified from prior ACTG ATI trials, including 4 PTCs who ini-

tiated ART during early HIV infection and 6 who initiated ART during chronic HIV infection (Supplemental Figure 1; supplemental material available online with this article; <https://doi.org/10.1172/JCI120549DS1>). The median duration of documented viral control was 63 weeks for the PTCs. To exclude continued ART use as the cause of HIV control, all PTCs underwent drug-level testing and no antiretroviral drugs were detectable during the treatment interruption. These 10 PTCs were compared with 16 NCs from the same studies and were well matched in their demographic characteristics (Table 1 and Supplemental Table 1). Nine individuals (4 PTCs

Table 1. Demographic characteristics of PTCs and NCs

	PTCs (N = 10)	NCs (N = 16)
Male, n (%)	8 (80)	11 (68.8)
Age, median, years	42	44
Early treated, n (%)	4 (40)	5 (31.3)
Years on ART, median	4.5	4.5
CD4 count, median	894	802
Race		
White, n (%)	7 (70)	10 (62.5)
Black, n (%)	1 (10)	3 (18.8)
Hispanic, n (%)	2 (20)	3 (18.8)

PTCs, posttreatment controllers; NCs, posttreatment noncontrollers; ART, antiretroviral therapy.

and 5 NCs) were treated during early infection as part of ACTG trial 371 (11), in which participants were treated with at least 52 weeks of ART prior to undergoing an ATI. The remaining individuals were treated during chronic infection. A median of 7.4 million peripheral blood mononuclear cells (PBMCs) were sampled from each participant from the pretreatment interruption time point, and 1,124 proviral genomes were obtained with a median of 48 proviral genomes per participant (Q1, Q3: 17, 58). In addition, we amplified 326 near-full-length proviral sequences from 7 PTCs at a posttreatment interruption time point. All proviral sequences were categorized as either intact or defective. The defects included internal inversions, hypermutations, deletions, premature stop codons, and defects in the packaging signal (*psi*) (Figure 1). For each participant, “genomes per million PBMCs” were calculated for total proviral genomes and for each proviral species.

In total, our data set was comprised of 137 intact and 1,313 defective proviral genomes. Prior to the treatment interruption, defective proviral genomes constituted the vast majority of the HIV reservoir (median 97% of the total proviral genomes), although substantial variation among individuals was observed (Q1, Q3: 87%, 99%) as demonstrated with 4 representative study participants (Figure 2A). To ensure that there was no cross-contamination between participant samples, we generated a neighbor joining tree including all the intact sequences obtained for the study and confirmed that all single-genome sequences clustered appropriately (Figure 2B).

Distinct proviral landscapes in PTCs and NCs. Until now, the intact proviral reservoir has not been described in PTCs. We hypothesized that a smaller intact proviral reservoir size before treatment interruption may be a determinant of posttreatment control. Our results showed that prior to treatment interruption, PTCs had approximately 7-fold lower levels of intact proviral genomes than NCs (IPGs, PTCs vs. NCs: median 0.04 vs. 0.28 copies/10⁶ PBMCs, $P < 0.05$, Figure 3A and Supplemental Figure 2). These intact proviral sequences likely represent the replication-competent reservoir, as we detected exact matches with plasma-derived sequences for a subset of participants (Supplemental Figure 3). We observed lower percentages of intact proviruses in the PTCs compared with NCs, although this difference did not reach statistical significance (median 1.4% vs. 4.1%, $P = 0.4$, Figure 3B).

Of note, there was a wide variation in the relative proportion of intact proviral genomes, including 8 participants whose viral reservoir contained more than 10% intact proviral genomes and 4 individuals with more than 20% intact proviral genomes.

Interestingly, levels of total proviral genomes (TPGs) were 7-fold lower in PTCs (median 1.6 vs. 11.1 copies/10⁶ PBMCs, $P < 0.001$, Figure 3C and Supplemental Figure 2) and were predominantly driven by levels of defective proviral genomes (DPGs, median 1.5 vs. 10.8 copies/10⁶ PBMCs, $P < 0.001$, Figure 3D and Supplemental Figure 2). There were no significant differences in the percentages of defective proviruses between PTCs and NCs (median 98.6% vs. 95.9%, Figure 3E). Among the readouts examined (IPGs, TPGs, and DPGs), levels of TPGs were the best reservoir marker to differentiate between PTCs and NCs, as 81% of NCs versus 0% of PTCs had TPGs greater than 4 copies/10⁶ PBMCs. No significant correlations between either pre-ATI CD4⁺ count or duration of ART treatment and any of the reservoir size measurements were detected. As a sensitivity analysis, we also calculated reservoir size per million CD4⁺ T cells, instead of PBMCs, and observed similar findings.

Next, we assessed whether certain types of defective proviruses were overrepresented in PTCs, as they can inform potential mechanisms underlying HIV control. Within the DPGs, PTCs had fewer proviral sequences with large deletions (median 1.1 vs. 10.5 copies/10⁶ PBMCs, $P < 0.001$) and fewer hypermutated proviral genomes (HPGs, median 0.2 vs. 1.2 copies/10⁶ PBMCs, $P < 0.01$, Figure 3F). However, the percentages of proviral sequences with large deletions or hypermutated proviral sequences were not significantly different between PTCs and NCs (Figure 3G). Furthermore, there was no difference between PTCs and NCs in the extent of apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBEC)-induced hypermutations relative to control (non-APOBEC) mutations (Figure 3H). The median proportions of each proviral species in the PTCs and NCs are shown in Figure 3I. In addition, we were interested in exploring whether treatment during early versus chronic infection changed the proportion of the different subsets of proviruses. In comparing the early and chronic-treated participants, chronic-treated NCs had higher levels of TPGs (median 15.7 vs. 2.6 copies/10⁶ PBMCs, $P < 0.05$) and DPGs (median 13.3 vs. 2.5 copies/10⁶ PBMCs, $P < 0.05$) compared with early treated NCs, but there were no significant differences in percentages of any proviral subset or in the copy numbers of IPGs or HPGs (Supplemental Figure 4).

To confirm that the proviral sequences we characterized as intact were replication competent, we isolated CD4⁺ T cells from 2 study participants, 1A8 (PTC) and 1C114 (NC), for whom large numbers of PBMCs were available. The cells were activated and cocultured with MOLT-4 cells in a viral outgrowth assay using a Transwell system. MOLT-4 cells from wells where virus was detected in the supernatant were subjected to near-full-length proviral amplification. Sequences obtained for each participant, 1A8 and 1C114, were nearly identical to the intact provirus, differing by only 4 and 5 nucleotides out of 9,020 nucleotides, respectively. These few nucleotide differences likely represent the expected variation arising during the 3-week culture period used to obtain each isolate (12, 13) and translated to 1 amino acid change in *env* for 1A8 and 2 changes in *env* for 1C114. Thus, we

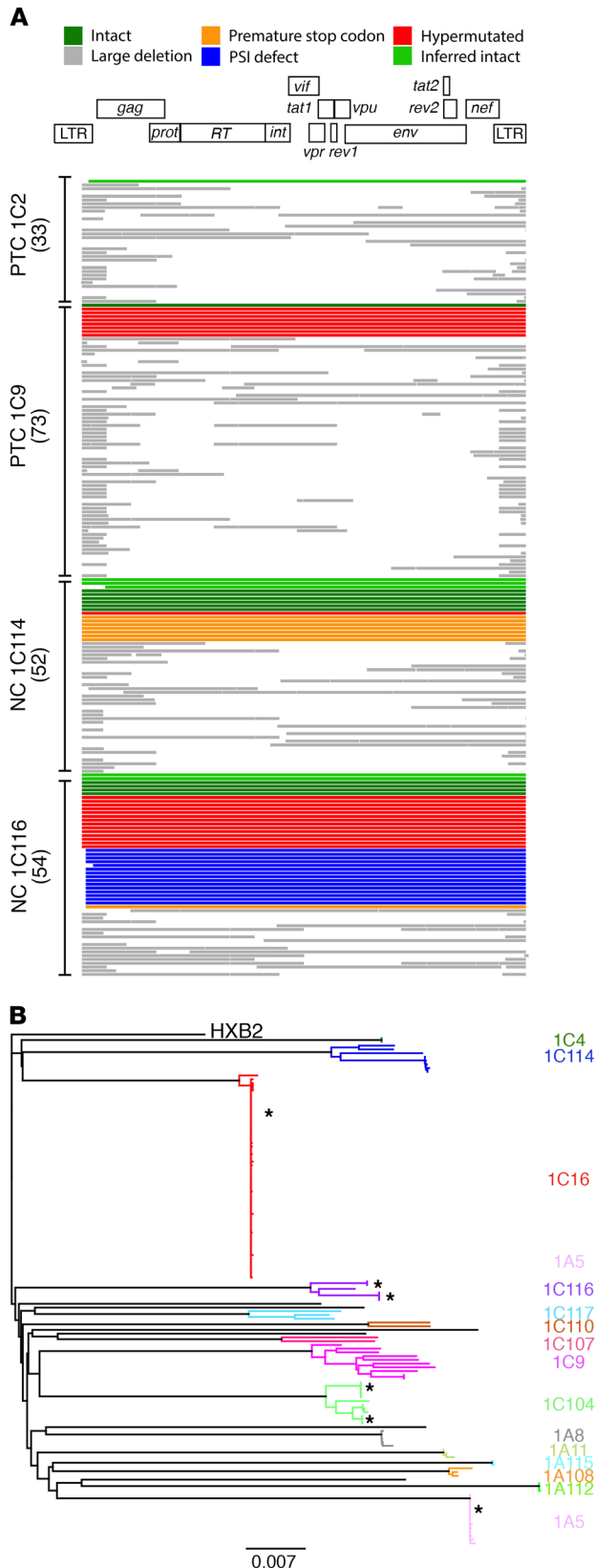


Figure 2. Near-full-length, single-template HIV-1 DNA amplicons. (A) Virograms reflecting the diversity of HIV DNA PCR products amplified from 4 representative participants (2 PTCs and 2 NCs, all treated during chronic infection). Numbers in parentheses indicate absolute frequency of analyzed sequences in each participant. **(B)** Combined neighbor-joining phylogenetic tree for all 137 intact proviral sequences showing no evidence of cross-participant sequence contamination, with HXB2 as an outgroup. Participants with more than 1 intact near-full-length proviral sequence are colored and those with only 1 identified intact sequence are depicted in black. The IDs of each participant are shown to the right of the tree. Asterisks denote clonal sequences detected more than once. PTC, posttreatment controllers; NC, posttreatment noncontrollers; PSI, packaging signal.

wanted to confirm that our bioinformatic calls of intact *env* in these proviral sequences reflect true functionality. So we separately tested the functionality of *env* of 4 different intact proviruses, including 1A8 and 1C114, by cotransfecting *env* PCR fragments with a Δenv -NL4-3 plasmid and determined the infectivity of the produced virus in a TZM-bl assay. Supplemental Figure 5 shows the infectivity titer for 4 recombinant viral constructs harboring intact proviral *env* from these participants, indicating that they are functional.

Stability of the intact HIV reservoir after treatment interruption in PTCs. Upon treatment interruption, levels of HIV DNA increase dramatically in most patients with HIV (14). We wanted to investigate whether this also occurred in PTCs after ATI, especially within the intact proviral reservoir. Levels of IPGs, TPGs, and DPGs were assayed for 7 PTCs with available samples at a late time point, a median of 71 weeks after treatment interruption. Unlike previously reported NCs (14), we detected no consistent increases in levels of IPGs, TPGs, and DPGs between the pre- and posttreatment interruption time points (Figure 3, J-L). The median ratio of pre/posttreatment interruption genomes per million PBMCs was 0.9 for IPGs, 1.3 for TPGs, and 1.4 for DPGs. No new HIV drug resistance mutations were detected in any of these individuals after treatment interruption compared with the preinterruption time point (15, 16).

No differences between PTCs and NCs in the clonal expansion of cells harboring intact proviral sequences. As illustrated in Figure 2B, our data set included several participants with a significant number of identical intact proviral sequences and likely represent clonally expanded sequences. Thus, we assessed whether the smaller intact proviral reservoir detected in PTCs is caused by differential rates of clonal expansion. To investigate this, we assessed the percentage of identical proviral sequences detected more than once and found no significant differences between PTCs and NCs in the proportion of clonally expanded proviral genomes. This was true regardless of whether the clonally expanded populations harbored intact (PTCs vs. NCs: median 0% in both cases) or defective proviral sequences (median 14% vs. 17%, Figure 4A and Supplemental Figure 6). Likewise, the absolute number of IPGs detected more than once was not significantly different between PTCs versus NCs, although PTCs did have a smaller number of DPGs detected more than once (median 0.2 vs. 2.7 copies/ 10^6 PBMCs, $P < 0.01$, Figure 4B). A sensitivity analysis of only participants who initiated ART during chronic infection yielded similar results. Of note, there were 4 participants who harbored a high percentage of clonal intact proviruses, 3 of whom initiated ART during chronic infection

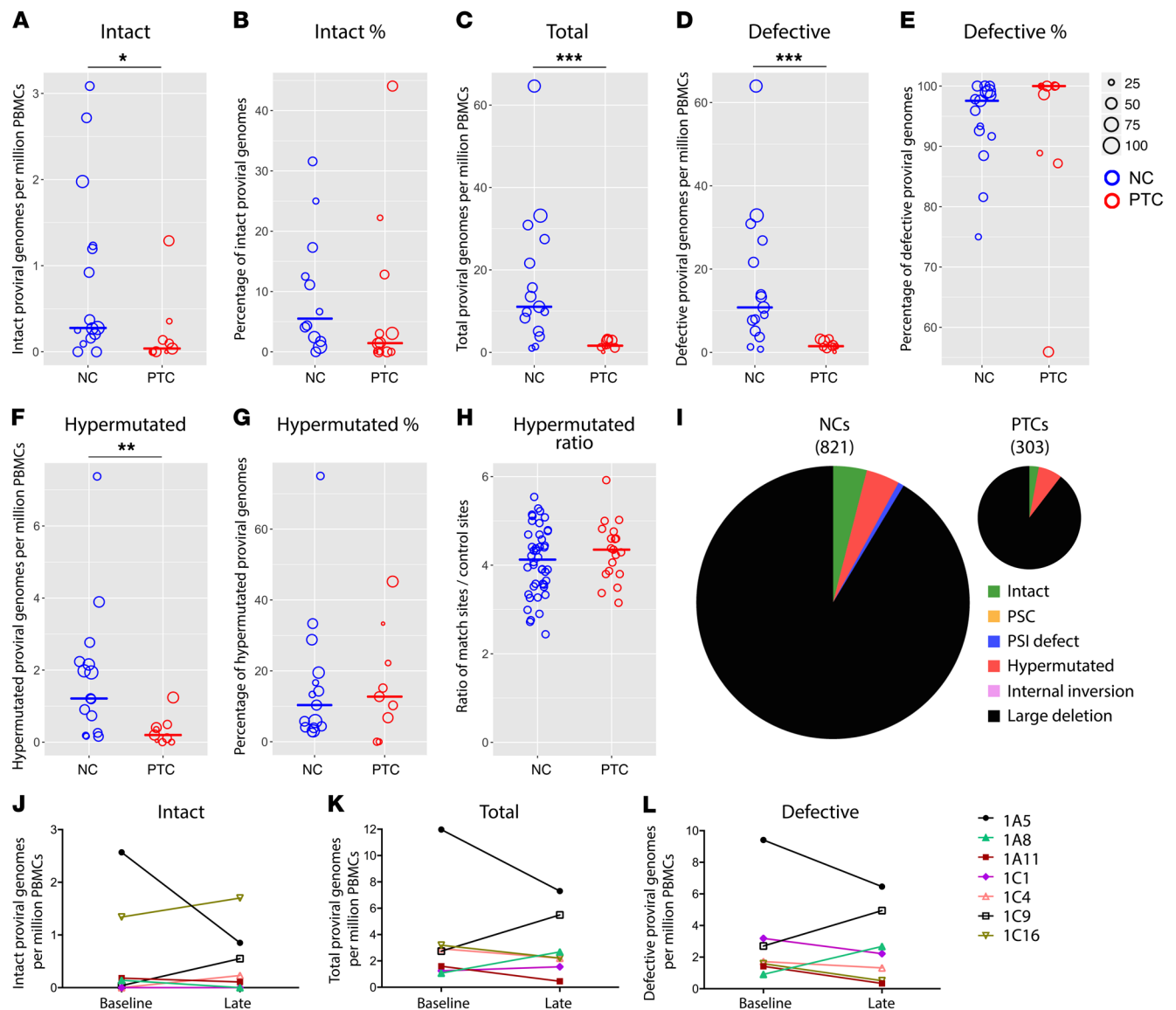


Figure 3. Comparison of reservoir measures between PTCs and NCs. (A) Copy numbers and (B) percentages of intact proviral genomes within total reservoir are compared between both groups. (C) Copy numbers of total proviral genomes per million PBMCs are compared between groups. (D) Copy numbers and (E) percentages of defective proviral genomes are depicted for each group. (F) Copy numbers and (G) percentages of hypermutated proviral genomes are depicted for each group. The size of each data point corresponds to total number of sequences amplified per participant as indicated in the legend in A–G. (H) A comparison between PTCs and NCs in the hypermutation rate ratio, calculated as the ratio of the number of match sites out of potential sites to the number of control sites out of potential sites. (I) Pie charts reflecting the median relative contribution of each proviral species for all participants, normalized to 100%. Numbers in parentheses indicate absolute frequency of analyzed sequences in each group. (J–L) Longitudinal analysis of intact (J), total (K), and defective (L) proviral genome copy numbers from the pre- and post-ATI time points for 7 PTCs. Each PTC is shown in a different color. Significance was calculated using a Wilcoxon rank-sum test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. NCs are depicted in blue and PTCs in red. PTC, posttreatment controllers; NC, posttreatment noncontrollers; PSI, packaging signal; PSC, premature stop codon.

(31% for patient 1C16, 26% for patient 1C104, and 9% for patient 1C116, Figure 4C) and 1 treated during early infection (21% for patient 1A5). These intact sequences could be grouped into a few clusters of identical sequences with the largest 2 clusters comprising at least 50% of all expanded intact sequences. Integration site analysis for these 4 participants support the near-full-length proviral sequencing results, showing that clonality estimates calculated by identical near-full-length proviral genomes were similar to those calculated from integration site analysis (Figure 4D). The frequency of clonally expanded sequences determined

by near-full-length sequencing was on average 75% (Q1, Q3: 61%, 82%) of the frequency determined by integration site analysis. Supplemental Table 2 contains a list of all integration sites obtained from these participants.

A possible driver of clonally expanded cells harboring intact proviruses could be T cell activation, given that it precipitates cellular proliferation and clonal expansion. However, we observed no significant correlation between percentages of HLA-DR⁺ CD38⁺ CD4⁺ or CD8⁺ T cells and either the proportion or absolute numbers of clonally expanded IPGs and DPGs.

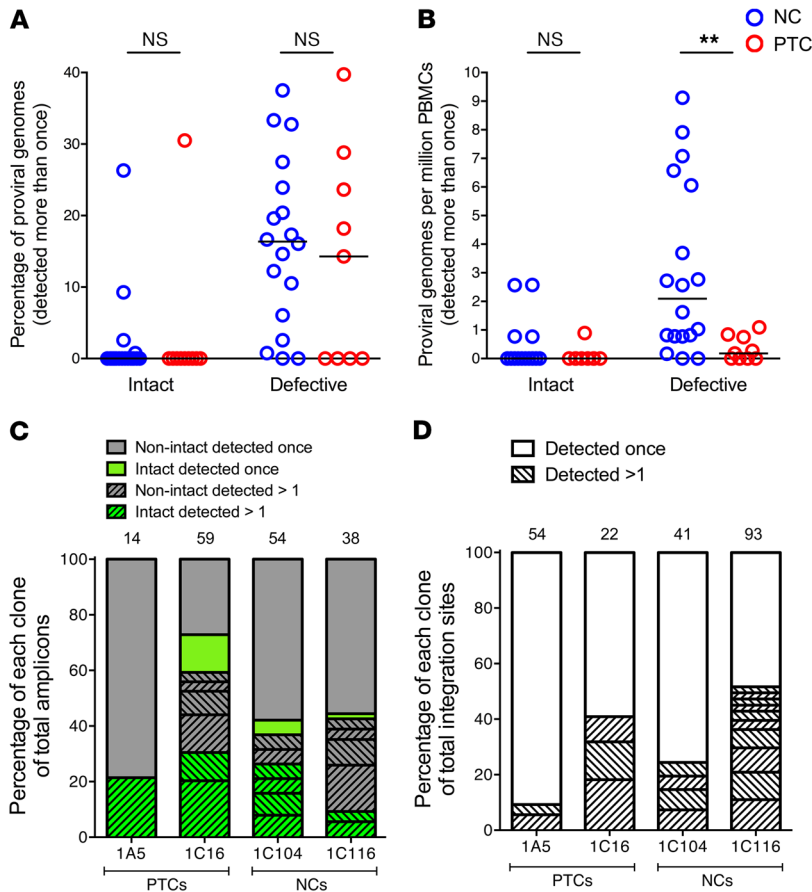


Figure 4. Contribution of clonal sequences to the proviral reservoir. (A) No difference between PTCs and NCs in percentages of either intact or defective proviral genome copies detected more than once. (B) PTCs have lower numbers of defective proviruses detected more than once, but no difference in intact copy numbers. Significance was calculated using a Wilcoxon rank-sum test, $**P < 0.01$. (C) Analysis of 4 participants who harbored a high percentage of intact proviruses detected more than once. Different stripe patterns depict different clones. Total number of amplicons is shown above each bar. (D) Integration site analysis of the same 4 participants in C. Different stripe patterns depict different clones. Total number of integration sites is shown above each bar. PTC, posttreatment controllers; NC, posttreatment noncontrollers.

HLA typing and HLA-associated escape polymorphisms. In order to examine the role of protective HLA alleles in posttreatment control, we analyzed available HLA typing for 6 of the PTCs and 11 of the NCs treated during chronic infection. Overall, protective HLA alleles were rare and found in only 1 participant of each group. We next surveyed for evidence of immune pressure on the proviral reservoir by examining inferred HLA-associated escape mutations within *gag*, *pol*, and *nef* in the context of each participant’s HLA class I profile (17). Inferred escape mutations matching the participants’ HLA allele(s) were present at a median of 25% of possible amino acid sites in *gag*, 17.4% in *pol*, and 25.8% in *nef* across all participants (Supplemental Figure 7, A–C). There were no significant differences between PTCs and NCs in the extent of mutations detected (Supplemental Figure 7, D–F). We also analyzed the proportion of inferred HLA escape mutations after treatment interruption in 4 PTCs and observed a modest increase in the percentage of HLA escape sites for at least 1 gene in 3 of the 4 individuals (Supplemental Figure 7, G–I). Trends in the longitudinal changes of the percentage of

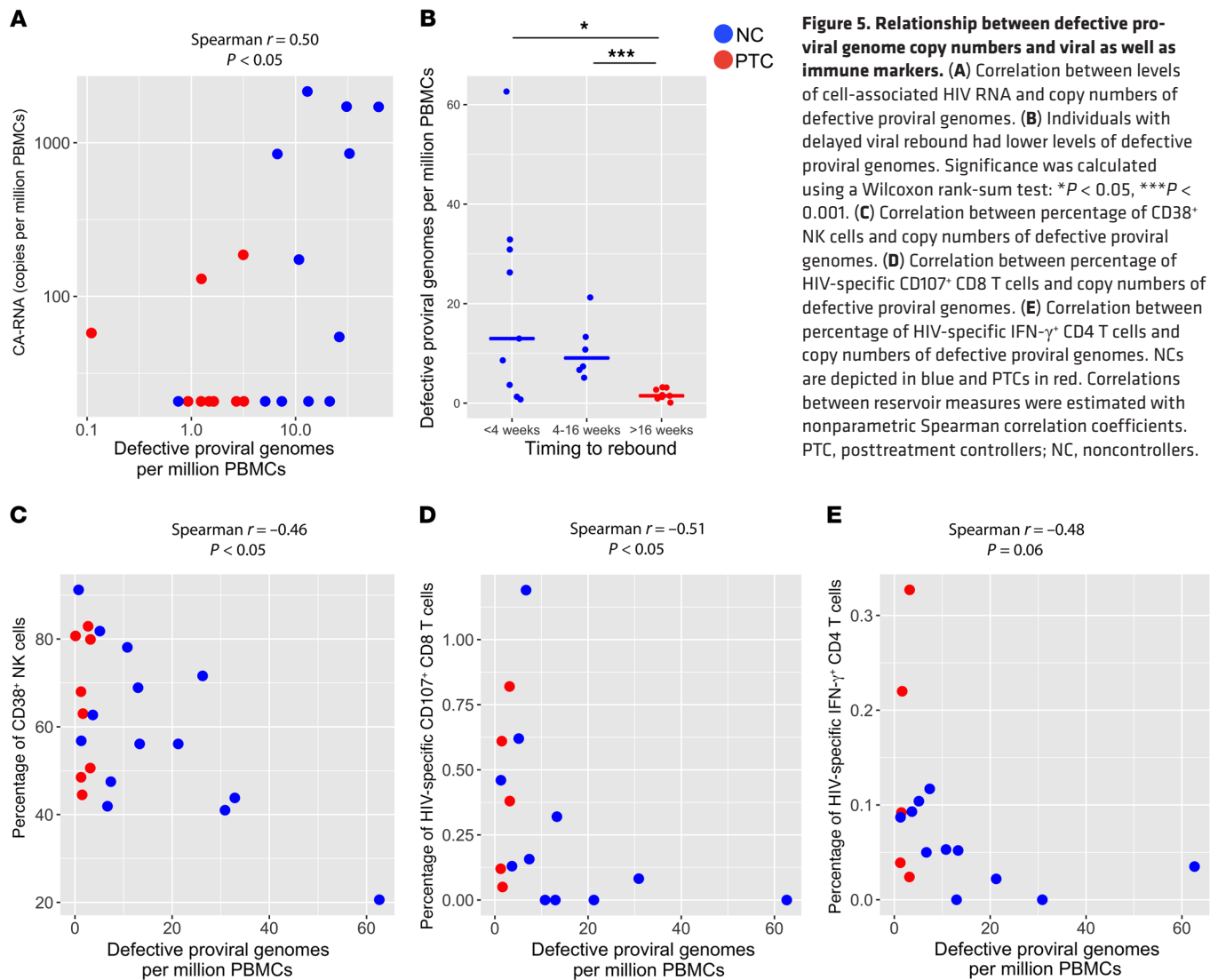
HLA escape sites for the intact proviruses were generally mirrored for the defective proviruses.

Defective proviral genome copies are associated with levels of CA-RNA and timing to viral rebound. Levels of intracellular HIV RNA expression have been shown to predict the timing of viral rebound after treatment interruption (1, 18). Thus, we assessed whether levels of intact provirus correlated with both levels of CA-RNA, as well as timing to viral rebound. Contrary to our expectations, the number of preinterruption IPGs was not significantly associated with unspliced CA-RNA levels as quantified by qPCR. However, the number of TPGs and DPGs were associated with levels of CA-RNA ($r = 0.50$, $P < 0.05$, Figure 5A and Supplemental Figure 8A). In addition, the number of TPGs and DPGs, but not IPGs, predicted the timing of subsequent viral rebound after ATI (Figure 5B and Supplemental Figure 8B). Individuals with late viral rebound (≥ 400 HIV-1 RNA copies/ml) were found to have the lowest TPG and DPG numbers (viral rebound >16 weeks vs. 4–16 weeks vs. <4 weeks = median DPGs 1.5 vs. 9.1 vs. 13.0 copies/ 10^6 PBMCs, respectively). There was no association among the levels of any specific type of defective sequences and either levels of CA-RNA or timing to viral rebound.

Association of defective proviral genomes with both innate and adaptive immune responses. Next, we hypothesized that the intact proviral reservoir size is dictated by levels of immune activation and the strength of the HIV-specific cellular immune response. Immune phenotyping was performed for $CD4^+$ / $CD8^+$ T cells and NK cells for markers of activation and cellular exhaustion prior to treatment interruption. The extent of NK cell function and HIV *gag*-specific T cell responses were assessed by intracellular

cytokine staining. Amongst all participants, none of the NK cell phenotypes or T cell responses were associated with levels of IPGs. However, higher $CD38^+$ NK cell percentages were associated with lower numbers of DPGs ($r = -0.46$, $P < 0.05$, Figure 5C). Similarly, higher levels of HIV-specific $CD8^+$ cells expressing $CD107a$ and HIV-specific IFN- γ -secreting $CD4^+$ cells were both associated with lower DPGs ($CD8^+$ $CD107a^+$: $r = -0.51$, $P < 0.05$; $CD4^+$ IFN- γ^+ : $r = -0.48$, $P = 0.06$; Figure 5, D and E). Given that the vast majority of the proviruses are defective, all the trends described are also true for TPGs (Supplemental Figure 8, C–E). There was no association among the levels of specific types of defect and any of the innate or adaptive immune responses. Furthermore, there were no significant associations between TPGs, DPGs, or IPGs and levels of soluble inflammatory markers D-dimer, IL-6, IP10, sCD14, sCD163, or CRP (Supplemental Figure 9).

Common motifs present at proviral deletion junctions. Given the high frequency of proviruses with large deletions, we assessed the presence of common motifs at the site of proviral deletions and



found that repeat elements were detected at the deletion junction in 11% of the proviral genomes with large deletions (Figure 6A). These repeat elements ranged in size from 3 to 20 nucleotides (Figure 6B). Interestingly, one specific repeat element residing at HXB2 coordinates 4781–4800 (TTTTAAAAGAAAAGGGGGGA) flanked the deletion junctions in 16 different proviral sequences originating from 8 study participants (2 PTCs and 6 NCs, Figure 6C). This element maps to both the polypurine tract located in *pol* (cPPT), as well as in *nef/U3* (3'PPT) and likely represents a hotspot for deletion. This sequence functions as a primer site for plus strand DNA synthesis and has been previously studied for its role in HIV replication (19–21).

Discussion

Up until now, the best-studied PTCs are the VISCONTI cohort (3, 22), but other isolated cases of PTCs have also been described (18, 23–32). In the VISCONTI study, 14 acutely treated participants were found to have levels of HIV DNA that were substantially lower than that of ART-treated individuals previously described in the literature (3). However, it is clear that the vast majority of

the HIV DNA reservoir is comprised of replication defective proviruses (7–10, 33) and an in-depth analysis of the proviral reservoir, including the frequency of intact proviruses, has yet to be performed for any of these PTCs. In this study, we provide the first detailed survey of the proviral genomic landscape in a PTC cohort and contrast it with that of NCs.

Our results show that prior to treatment interruption, the DNA reservoir size, defined as either total or intact proviral genomes, was a median 7-fold lower in PTCs than in NCs. Specifically, 81% of NCs had greater than 4 total proviral genome copies/ 10^6 PBMCs, whereas none of the PTCs had a total reservoir size that reached this threshold. Low levels of total HIV DNA have been previously associated with short-term delays in viral rebound timing (29, 34), as exemplified by the report of 2 Boston patients who underwent allogeneic hematopoietic stem cell transplantation with susceptible donor cells. In both patients, proviral DNA was reduced to undetectable levels after transplantation and viral rebound was delayed until 12 and 32 weeks after ATI despite an immune system that was functionally naive to HIV (35). This observation led some to predict that at least a 3- to 4- \log_{10} reduction in the HIV reservoir

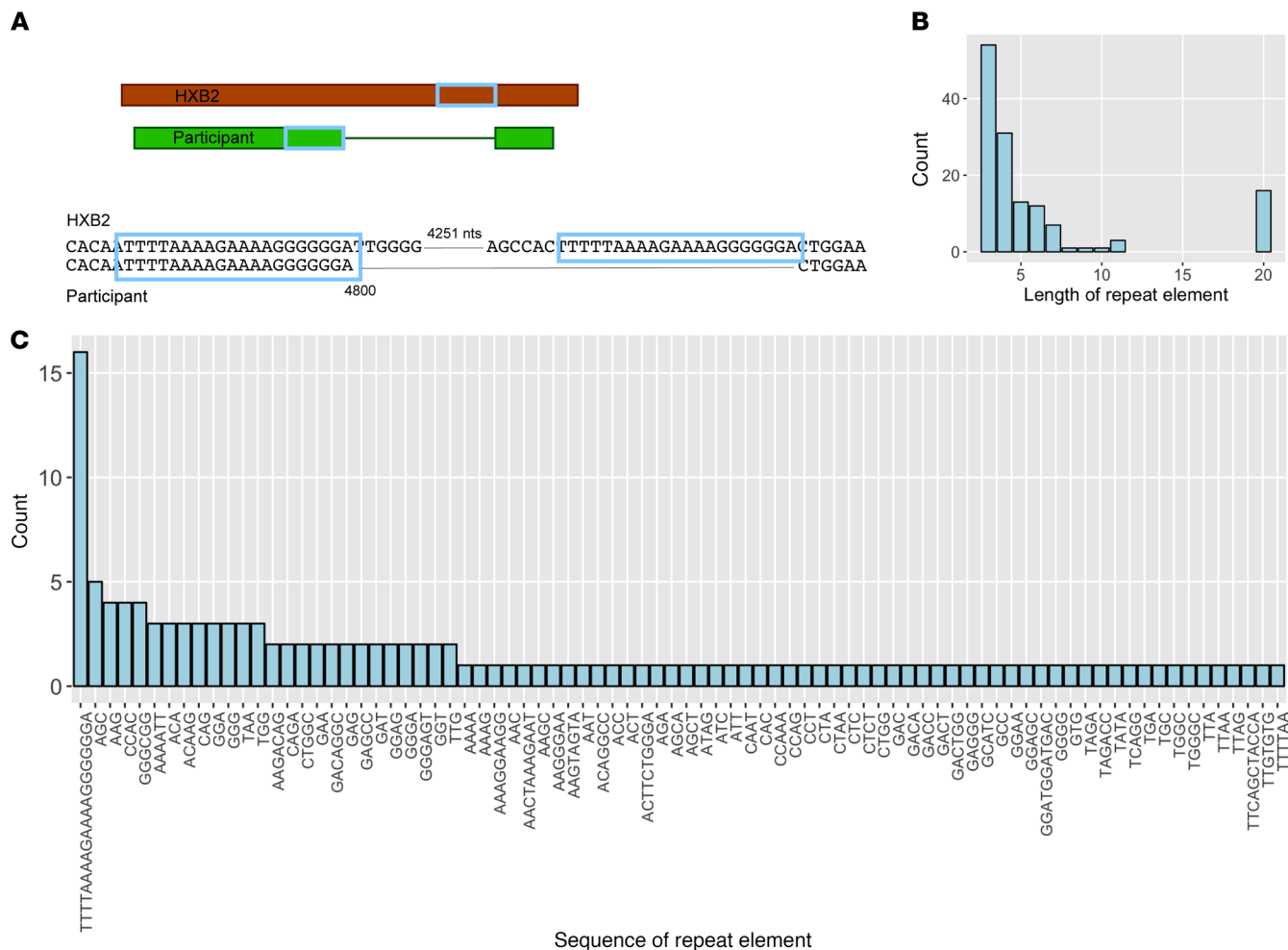


Figure 6. Detection of repeat elements at HIV proviral deletion junctions. (A) Illustration of a common pattern present flanking both ends of the deletion junction. **(B)** Histogram of number of proviral genomes detected with repeat elements of varying lengths. One specific 20 base nucleotide repeat is overrepresented. **(C)** Histogram of the number of times each repeat was detected. nts, nucleotides.

may be needed to achieve sustained HIV remission (36, 37). Our finding of a substantial difference in time to rebound between PTCs and NCs despite relatively modest (7-fold) differences in total and intact proviral genomes suggests that factors in addition to reservoir size must play an important role in determining time to rebound. Such factors could include differences in viral replication rate and/or robust anti-HIV immune responses. This is underscored by the finding that despite prolonged treatment interruption, levels of total and intact proviral reservoir did not significantly increase in PTCs, similar to results from the VISCONTI study. This finding contrasts with the pattern reported for other HIV-infected subjects, where the HIV DNA burden markedly increases upon treatment interruption due to viral replication (14, 38, 39).

Given the rarity of PTCs and safety concerns involving treatment interruption studies, future HIV cure trials would benefit from a set of biomarkers that could predict which participants are more likely to be PTCs prior to treatment interruption. Our finding that the majority of NCs, but none of the PTCs, had greater than 4 total proviral genomes/10⁶ PBMCs before treatment interruption suggests that quantification of total proviral genome copy numbers

could contribute to the prediction of sustained HIV remission. PTCs had a greater variation in the number of intact proviral genomes that hindered its use in differentiating between the 2 groups. In particular, more than 40% of the total proviral reservoir for PTC patient 1C16 was intact, including 2 large clonal populations. This observation highlights that the role of the intact proviral reservoir size in predicting the timing of viral rebound remains unclear, which may be due to its low frequency within the proviral landscape. Additional analysis of transcriptional activity, integration sites, and viral replication fitness of these intact proviruses is needed.

Similar to previously published reports (7–10, 33), we found that the vast majority of proviral genomes amplified from participant-derived DNA were defective and only a small proportion had intact open reading frames. This was true both for participants treated in either the early or the chronic stage of HIV infection. We were intrigued to find that for PTCs, the proportion of intact proviruses was almost 3 times lower (1.4%) than in NCs (4.1%), although this comparison did not reach statistical significance and the characterization of additional participants is needed to further explore this finding. We also found no sig-

nificant differences in the frequency of any particular defective proviral species between PTCs and NCs. Of note, our analysis of defective proviral genomes revealed the proportionally higher prevalence of a specific proviral deletion, represented in 8 distinct participants and reported previously in 2 other cohorts (8, 9). This pattern comprises a 20 nucleotide repeat sequence flanking the deletion junction, representative of the polypurine tract, and appears to be a hotspot for proviral deletions.

HIV integration site analysis has suggested that clonal proliferation of HIV-infected cells may play an important role in the proliferation and persistence of HIV-infected cells (40, 41). Other groups have shown that clonally expanded populations of HIV-infected cells can harbor intact, replication-competent proviruses (12, 33, 42, 43). A recent report provided definitive evidence that latent cells that can be reactivated can arise by clonal expansion *in vivo* by determining that the cells harbored identical TCR sequences through scRNA-seq (44). Thus, we evaluated whether clonal expansion of intact proviruses could be playing a role in the reservoir size differences between NCs and PTCs by assessing the number and frequency of identical proviral sequences detected more than once. We observed no difference between PTCs and NCs in the proportion of identical intact proviral sequences detected more than once. This suggests that the clonal expansion rate of cells harboring intact proviruses is similar between PTCs and NCs and therefore does not explain the smaller reservoir size in PTCs or their ability to maintain viral suppression.

We also evaluated the relationship between the proviral landscape and both levels of viral transcription and immune activity. The level of unspliced CA-RNA during suppressive ART has been shown in several studies to be predictive of the timing of viral rebound after ART discontinuation (1, 18), which suggested that CA-RNA may largely reflect transcription of the intact provirus. Surprisingly, we found that the level of total and defective but not intact proviral genomes was significantly associated with CA-RNA levels (Figure 5A and Supplemental Figure 8A). These results reinforce a recent finding that defective proviruses express intracellular HIV RNA (45) and that the bulk of the CA-RNA quantified is likely defective given that the vast majority of proviruses are replication incompetent (8, 9, 46). Any relationship between the size of the intact proviral reservoir and CA-RNA levels is likely obscured by the overwhelming number of defective proviruses also expressing HIV RNA.

In addition, we found an association between higher levels of NK cell activation and HIV-specific T cell activity with lower levels of total and defective proviral genomes. This suggests that a more robust HIV-specific immune response can identify and effectively eliminate HIV-infected cells, even those harboring defective proviruses. Combined with the CA-RNA results, these findings support the concept that defective HIV genomes can lead to viral RNA transcription and antigen production (8, 46), which leads to interactions of cells harboring defective HIV genomes with both the innate and adaptive immune system. Further studies are needed to explore this finding.

We also found that the level of defective but not intact proviral genomes was associated with the timing of viral rebound after ART interruption. The reason behind this counterintuitive result is unclear, but possible explanations include: (a) the size of the defective proviral genome within the peripheral blood reflects the

overall true reservoir size within the body, including within anatomic sites, and (b) given the low frequency of intact proviruses, it is difficult to detect a relationship with timing of viral rebound and a larger sample may be needed.

The limitations of this study should also be acknowledged. First, there were individuals with extremely small reservoir sizes and relatively few amplified proviral sequences. This may affect the accuracy of the estimated frequency of different proviral species. However, the results were largely unchanged if the analysis was restricted to only participants with more than 20 proviral sequences. Second, *in vitro* reconstitution of the reported full-length intact proviruses is needed to confirm that they are indeed replication competent (9), although we were able to detect identical matches between intact proviral and plasma-derived sequences and obtained highly similar sequences in viral outgrowth assays.

In this report, we provide the most comprehensive survey to date of the proviral landscape in both PTCs and NCs prior to treatment interruption. The results highlight the restricted intact and defective HIV reservoir in PTCs and suggest that total proviral genome burden could act as the first biomarker for identifying PTCs. Additional studies are needed to confirm these findings and to explore whether inefficient viral replication and/or robust anti-HIV immune responses are also present in these individuals. We also found that the size of the defective proviral reservoir was associated with levels of CA-RNA and several NK and T cell markers. Until recently, the defective proviral reservoir appears to have been underappreciated, but these results support the concept that defective proviruses are far from quiescent (46) and are likely playing key roles in immune activation, shaping anti-HIV immune responses, and may be a marker of the timing of plasma viral emergence after ART discontinuation.

Methods

Study population and samples. PTCs were identified from several ACTG analytic treatment interruption studies (11, 47–50). PTCs were defined as individuals who were on suppressive ART and after ATI, maintained viral loads of no more than 400 HIV RNA copies/ml for at least 24 weeks (short-term [≤ 2] viral loads > 400 HIV RNA copies/ml were not exclusionary). Post-ATI plasma samples were tested for ART levels and drugs were not detected. NCs were individuals with available stored samples who did not meet the PTC criteria and selected to match the same study arms. All participants had PBMC samples available immediately before ATI (i.e., baseline) for near-full-length proviral sequencing. For a subset of participants with available samples, proviral sequencing was also performed on a post-ATI time point collected a median of 71 weeks (Q1, Q3: 45 weeks, 84 weeks) after ART discontinuation.

Specifically, early treated participants were stratified by predetermined criteria in the original study protocol to either the acute infection group or the recent infection group (11). Acute infection was defined as having a plasma HIV RNA concentration more than 2,000 copies/ml within 14 days of study entry and either a negative ELISA, or a positive ELISA but a negative or indeterminate Western blot, or a positive ELISA and Western blot in conjunction with either a negative ELISA or a plasma HIV RNA concentration less than 2,000 copies/ml in the 30 days prior to entry. The general intention of these criteria was to identify cases in which HIV infection had occurred within the 4

weeks prior to study entry. Recent infection was defined as a positive ELISA and Western blot within the 14 days prior to entry but a negative ELISA or plasma HIV RNA concentration less than 2,000 copies/ml within the 31 to 90 days before entry or a positive ELISA and Western blot and a nonreactive detuned ELISA in patients with more than 200 CD4⁺ cells/ μ l all within the 21 days before study entry.

Proviral sequencing. Next-generation single-genome sequencing (NG-SGS) utilizes the Illumina deep sequencing platform to efficiently sequence the single-genome, near-full-length proviral amplicons. Single-genome sequencing allows the isolation of full-length proviral HIV without the confounding concerns of PCR recombination frequently seen in bulk PCR products. Limiting dilution proviral amplification was performed using a technique adapted from a previously published protocol (51). DNA was extracted from cryopreserved PBMCs using the QIAmp DNA Mini Kit (Qiagen) and used at limiting dilution for nested PCR amplification with Platinum Taq HiFi polymerase (Thermo Fisher Scientific). The sequences of the primers for the first- and second-round PCR reactions were as follows: first-round forward primer: 5'-AAATCTCTAGCAGTGGCGCCCGAACAG-3', first-round reverse primer: 5'-TGAGG-GATCTCTAGTTACCAGAGTC-3'; second-round forward primer: 5'-GCGCCCGAACAGGGACYTGAAARCGAAAG-3', second-round reverse primer: 5'-GCACTCAAGGCAAGCTTTATTGAGGCTTA-3'. These primers correspond to HXB2 coordinates 623–649, 9662–9686, 638–666, and 9604–9632, respectively. PCR amplicons were sheared and Illumina barcoded libraries were constructed and pooled. Sequencing was performed on the Illumina MiSeq platform and amplicons were assembled using the UltraCycler v1.0. Automated de novo sequence assembly generated a continuous fragment of HIV-1 proviral DNA, which was fed into an automated in-house pipeline to determine proviral genome intactness (52), as previously published (33). Briefly, the sequences were aligned to HXB2 to identify sequence defects (e.g., internal deletions, premature stop codons, out-of-frame mutations, internal inversions, and packaging signal defects). The sequences were then run in the Los Alamos HIV Sequence Database Hypermut program to identify hypermutated sequences (53). Proviral sequences that lacked any of the above-mentioned defects were classified as intact. Near-full-length assembled sequences missing one of the PCR primer binding sites, but otherwise meeting the criteria for an intact provirus, were labeled “inferred intact.” A clonal cluster analysis was also performed to detect genomes that were 100% identical across the entirety of their assembled length. The unit “genomes per million PBMCs” was calculated for total proviral genomes and for each proviral species based on the quantity of DNA assayed.

Additionally, proviral sequences were analyzed in the context of each participant's HLA class I profile for the presence of known HLA-associated polymorphisms in HIV-1 subtype B as defined in a published reference list (17). We also assessed the presence of repeat elements flanking the deletion junctions using a Python script.

Plasma viral sequencing. Plasma viral RNA was extracted using the QIAamp Viral RNA Mini Kits and SGS of HIV-1 *pro-rt* (HXB2 coordinates 2853–3869) performed as previously published (54, 55). Using ClustalW, the resulting plasma-derived single-genome sequences were aligned with proviral sequences harboring the *pro-rt* region.

Quantification of HIV DNA and CA-RNA. CA-RNA was isolated from cryopreserved PBMCs using the AllPrep DNA/RNA Mini Kit

(Qiagen). Unspliced CA-RNA levels were quantified using a real-time PCR approach with primers/probes targeting conserved regions of HIV LTR/*gag* as previously described (1). Cell numbers were quantified by qPCR measurements of CCR5 DNA copy numbers. Measurement of IPO-8 transcripts in total RNA was used as an internal quality control to assess the efficiency of RNA extraction (56).

Immune phenotyping and intracellular cytokine staining. Approximately 10⁶ PBMCs each were used for T and NK cell phenotyping. T cells were stained with blue viability dye (Invitrogen) at 37°C for 20 minutes and then followed with antibodies targeting CD3, CD8, CD4, CD45RO, CD95, HLA-DR, CD38, and PD-1 at 4°C for 20 minutes. NK cells were stained with blue viability dye and antibodies targeting CD3, CD19, CD16, CD56, CD69, CD38, CD57, NKG2D, NKp30, and NKp46.

T cell intracellular cytokine staining (ICS) was performed on PBMCs stimulated with HIV *gag* peptide pool and NK cell ICS was performed with PBMCs stimulated with K562 cells. For T cells, approximately 10⁶ cells were stimulated overnight with anti-CD28/49d (0.5 μ g/ml; BD Biosciences) and 2 μ g/ml of synthetic peptides (overlapping 15- to 20-mer *gag* peptide pools spanning the entire clade B consensus sequence of the HIV-1 *gag* sequence). NK cells were stimulated with K562 cells with an effector-to-target ratio of 10:1. Brefeldin A (1 μ g/ml; BioLegend), Monesin Solution (1 μ g/ml; BioLegend), and CD107a antibody were added after a 1-hour incubation and cultured for an additional 5 hours. Cells stimulated with PMA (2.5 μ g/ml) and ionomycin (0.5 μ g/ml) served as a positive control and R10 medium alone served as a negative control. After stimulation, the cells were stained with surface antibodies against CD3, CD19, CD16, CD56, and blue viability dye at 4°C for 20 minutes. Subsequently, cells were treated with a fixation and permeabilization solution per the manufacturer's protocol. Cells were stained for 20 minutes at room temperature with antibodies directed to IFN- γ , IL-2, CD107a, and TNF- α . Cells were then fixed by 2% PFA (Affymetrix), acquired on an LSR Fortessa flow cytometer (BD Biosciences), and analyzed using FlowJo software (version v10, Tree Star). The proportion of cytokine-secreting cells had to be greater than 0.1% after subtraction to be considered as a positive response.

Soluble markers of inflammation. Soluble markers of inflammation were measured by ELISA. Plasma from all participants was analyzed for levels of IL-6 (HS600B, R&D Systems), sCD14 (DC140, R&D Systems), IFN- γ , IP10 (DIP100, R&D Systems), sTNFR-I and sTNFR-II (DRT100/200, R&D Systems), and D-dimer (Diagnostica Stago) per the manufacturers' protocols.

Infectivity of recombinant viruses. We generated recombinant viruses encoding patient-derived *env* sequences by cotransfecting 293T cells with Δenv -NL4-3 plasmid and *env* PCR fragments. Virus was harvested and propagated in U87-CCR5 and U87-CXCR4 cells for 7–10 days. Afterwards, we tested the infectivity of the virus in a TZM-bl infectivity assay. TZM-bl cells (NIH AIDS Reagent Program) are a permissive HeLa cell clone that contains a *Tat*-regulated reporter gene for β -galactosidase under the control of the HIV-1 LTR.

Integration site analysis. Integration sites were determined using the integration site loop amplification (ISLA) technique on genomic DNA isolated from PBMCs before treatment interruption, as previously published (41).

Viral outgrowth assays. CD4⁺ T cells were isolated by negative selection and stimulated with PHA (2 μ g/ml), rhIL-2 (50 U/ml), and

irradiated allogeneic PBMCs from an HIV-negative donor in the lower compartment of a Transwell. On day 14 of culture, the supernatant from each well was tested in a TZM-bl infectivity assay. MOLT-4 cells were added to the upper insert of Transwells with detectable virus on day 16 and harvested for DNA extraction on day 20, followed by near-full-length proviral amplification and sequencing.

Statistics. Statistical analysis was performed with SAS Studio (Release 3.6, University Edition) and Prism (v7, GraphPad). Wilcoxon rank-sum tests were used to compare reservoir measures between groups. Correlations between reservoir measures and other viral/immune markers were estimated with nonparametric Spearman's correlation coefficients. $P < 0.05$ was considered significant. Data were summarized as individual data plots with lines depicting median values. Where indicated, the size of each data point corresponds to the total number of proviral genomes sequenced for that participant.

Study approval. Participant samples were collected according to protocols approved by the respective institutional review boards. Study participants gave written informed consent in accordance with the Declaration of Helsinki.

Author contributions

RS, RTG, ML, XGY, DRK, and JZL developed the concept and design. RS, BE, LMA, ZH, YW, GN, MML, EPA, and XS performed the experiments. RS, GQL, CG, ZLB, ML, RJB, EA, and JZL were involved in the bioinformatic and statistical analysis. MJM, DS,

DMM, RM, PV, and EC contributed samples. All authors were involved in the writing of the manuscript.

Acknowledgments

This work was supported in part by National Institutes of Health/National Institute of Allergy and Infectious Diseases grants: AI125109 (to JZL), the Harvard University Center for AIDS Research (5P30AI060354-08 to JZL and RTG, 5P30AI060354-14 to GQL), UM1AI068634 (Statistical and Data Management Center of the AIDS Clinical Trials Group), UM1AI068636 (AIDS Clinical Trials Group), a subcontract from UM1AI106701 to the Harvard Virology Support Laboratory (to JZL), UM1AI126617 (to ZLB and EC), and UM1AI069423 (to DM). ZLB is supported by a Scholar Award from the Michael Smith Foundation for Health Research. We are grateful for the contributions of the participants who made this study possible. We thank the staff and principal investigators of ACTG studies A371, A5024, A5068, A5170, and A5197. We appreciate the support of Nicole Stange-Thomann and the staff of the MGH sequencing core facility. We thank Wei-Shau Hu, Tsibris, and Kuritzkes labs for valuable feedback.

Address correspondence to: Jonathan Z. Li, Brigham and Women's Hospital, 65 Landsdowne Street, Room 421, Cambridge, Massachusetts 02139, USA. Phone: 617.768.8476; Email: jli@bwh.harvard.edu.

- Li JZ, et al. The size of the expressed HIV reservoir predicts timing of viral rebound after treatment interruption. *AIDS*. 2016;30(3):343-353.
- Hütter G, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med*. 2009;360(7):692-698.
- Sáez-Cirión A, et al. Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS Pathog*. 2013;9(3):e1003211.
- Etemad B, et al. Viral and immune characteristics of HIV post-treatment controllers in ACTG studies. Presented at: CROI; Boston, Massachusetts, USA; February 22-25, 2016. Abstract 347.
- Conway JM, Perelson AS. Post-treatment control of HIV infection. *Proc Natl Acad Sci U S A*. 2015;112(17):5467-5472.
- Cockerham LR, Hatano H, Deeks SG. Post-treatment controllers: role in HIV "Cure" research. *Curr HIV/AIDS Rep*. 2016;13(1):1-9.
- Bruner KM, et al. Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat Med*. 2016;22(9):1043-1049.
- Imamichi H, et al. Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. *Proc Natl Acad Sci U S A*. 2016;113(31):8783-8788.
- Ho YC, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*. 2013;155(3):540-551.
- Hiener B, et al. Identification of genetically intact HIV-1 proviruses in specific CD4⁺ T cells from effectively treated participants. *Cell Rep*. 2017;21(3):813-822.
- Volberding P, et al. Antiretroviral therapy in acute and recent HIV infection: a prospective multicenter stratified trial of intentionally interrupted treatment. *AIDS*. 2009;23(15):1987-1995.
- Hosmane NN, et al. Proliferation of latently infected CD4⁺ T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J Exp Med*. 2017;214(4):959-972.
- Mansky LM, Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol*. 1995;69(8):5087-5094.
- Clarridge KE, et al. Effect of analytical treatment interruption and reinitiation of antiretroviral therapy on HIV reservoirs and immunologic parameters in infected individuals. *PLoS Pathog*. 2018;14(1):e1006792.
- Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res*. 2003;31(1):298-303.
- Shafer RW. Rationale and uses of a public HIV drug-resistance database. *J Infect Dis*. 2006;194(suppl 1):S51-S58.
- Carlson JM, et al. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. *J Virol*. 2012;86(24):13202-13216.
- Sneller MC, et al. A randomized controlled safety/efficacy trial of therapeutic vaccination in HIV-infected individuals who initiated antiretroviral therapy early in infection. *Sci Transl Med*. 2017;9(419):eaan8848.
- Miles LR, Agresta BE, Khan MB, Tang S, Levin JG, Powell MD. Effect of polypurine tract (PPT) mutations on human immunodeficiency virus type 1 replication: a virus with a completely randomized PPT retains low infectivity. *J Virol*. 2005;79(11):6859-6867.
- Charneau P, Alizon M, Clavel F. A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *J Virol*. 1992;66(5):2814-2820.
- Schultz SJ, Zhang M, Kelleher CD, Champoux JJ. Analysis of plus-strand primer selection, removal, and reutilization by retroviral reverse transcriptases. *J Biol Chem*. 2000;275(41):32299-32309.
- Hocqueloux L, et al. Long-term immunovirologic control following antiretroviral therapy interruption in patients treated at the time of primary HIV-1 infection. *AIDS*. 2010;24(10):1598-1601.
- Maenza J, et al. How often does treatment of primary HIV lead to post-treatment control? *Antivir Ther (Lond)*. 2015;20(8):855-863.
- Stöhr W, et al. Duration of HIV-1 viral suppression on cessation of antiretroviral therapy in primary infection correlates with time on therapy. *PLoS ONE*. 2013;8(10):e78287.
- Lodi S, et al. Immunovirologic control 24 months after interruption of antiretroviral therapy initiated close to HIV seroconversion. *Arch Intern Med*. 2012;172(16):1252-1255.
- Perkins M, et al. Post-treatment control of HIV infection in an early diagnosed well characterized military cohort of chronically HIV-1-infected subjects. *Open Forum Infect Dis*. 2015;2(suppl 1):1072.
- Goujard C, et al. HIV-1 control after transient antiretroviral treatment initiated in primary infection: role of patient characteristics

- and effect of therapy. *Antivir Ther (Lond)*. 2012;17(6):1001-1009.
28. Salgado M, et al. Prolonged control of replication-competent dual-tropic human immunodeficiency virus-1 following cessation of highly active antiretroviral therapy. *Retrovirology*. 2011;8:97.
 29. Assoumou L, et al. A low HIV-DNA level in peripheral blood mononuclear cells at antiretroviral treatment interruption predicts a higher probability of maintaining viral control. *AIDS*. 2015;29(15):2003-2007.
 30. Maggiolo F, Di Filippo E, Comi L, Callegaro A. Post-treatment controllers after treatment interruption in chronically HIV-1-infected patients. *AIDS*. 2018;32(5):623-628.
 31. Van Gulck E, et al. Immune and viral correlates of "secondary viral control" after treatment interruption in chronically HIV-1 infected patients. *PLoS One*. 2012;7(5):e37792.
 32. Frange P, et al. HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: a case report. *Lancet HIV*. 2016;3(1):e49-e54.
 33. Lee GQ, et al. Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4⁺ T cells. *J Clin Invest*. 2017;127(7):2689-2696.
 34. Williams JP, et al. HIV-1 DNA predicts disease progression and post-treatment virological control. *Elife*. 2014;3:e03821.
 35. Henrich TJ, et al. Antiretroviral-free HIV-1 remission and viral rebound after allogeneic stem cell transplantation: report of 2 cases. *Ann Intern Med*. 2014;161(5):319-327.
 36. Hill AL, Rosenbloom DI, Fu F, Nowak MA, Siliciano RF. Predicting the outcomes of treatment to eradicate the latent reservoir for HIV-1. *Proc Natl Acad Sci U S A*. 2014;111(37):13475-13480.
 37. Hill AL, et al. Real-time predictions of reservoir size and rebound time during antiretroviral therapy interruption trials for HIV. *PLoS Pathog*. 2016;12(4):e1005535.
 38. Garbuglia AR, et al. HIV-1 DNA burden dynamics in CD4 T cells and monocytes in patients undergoing a transient therapy interruption. *J Med Virol*. 2004;74(3):373-381.
 39. Calin R, et al. Treatment interruption in chronically HIV-infected patients with an ultralow HIV reservoir. *AIDS*. 2016;30(5):761-769.
 40. 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.
 41. 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.
 42. Bui JK, et al. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS Pathog*. 2017;13(3):e1006283.
 43. 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.
 44. Cohn LB, et al. Clonal CD4⁺ T cells in the HIV-1 latent reservoir display a distinct gene profile upon reactivation. *Nat Med*. 2018;24(5):604-609.
 45. 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.
 46. Pollack RA, et al. Defective HIV-1 Proviruses are expressed and can be recognized by cytotoxic T lymphocytes, which shape the proviral landscape. *Cell Host Microbe*. 2017;21(4):494-506.e4.
 47. Schooley RT, et al. AIDS clinical trials group 5197: a placebo-controlled trial of immunization of HIV-1-infected persons with a replication-deficient adenovirus type 5 vaccine expressing the HIV-1 core protein. *J Infect Dis*. 2010;202(5):705-716.
 48. Kiest DJ, et al. Interruption of antiretroviral treatment in HIV-infected patients with preserved immune function is associated with a low rate of clinical progression: a prospective study by AIDS Clinical Trials Group 5170. *J Infect Dis*. 2007;195(10):1426-1436.
 49. Kilby JM, et al. A randomized, partially blinded phase 2 trial of antiretroviral therapy, HIV-specific immunizations, and interleukin-2 cycles to promote efficient control of viral replication (ACTG A5024). *J Infect Dis*. 2006;194(12):1672-1676.
 50. Jacobson JM, et al. Evidence that intermittent structured treatment interruption, but not immunization with ALVAC-HIV vCP1452, promotes host control of HIV replication: the results of AIDS Clinical Trials Group 5068. *J Infect Dis*. 2006;194(5):623-632.
 51. Li B, et al. Rapid reversion of sequence polymorphisms dominates early human immunodeficiency virus type 1 evolution. *J Virol*. 2007;81(1):193-201.
 52. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria. <http://www.R-project.org/>. Accessed July 19, 2018.
 53. Rose PP, Korber BT. Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. *Bioinformatics*. 2000;16(4):400-401.
 54. 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.
 55. Kearney M, et al. Frequent polymorphism at drug resistance sites in HIV-1 protease and reverse transcriptase. *AIDS*. 2008;22(4):497-501.
 56. Ledderose C, Heyn J, Limbeck E, Kreth S. Selection of reliable reference genes for quantitative real-time PCR in human T cells and neutrophils. *BMC Res Notes*. 2011;4:427.