JCI The Journal of Clinical Investigation

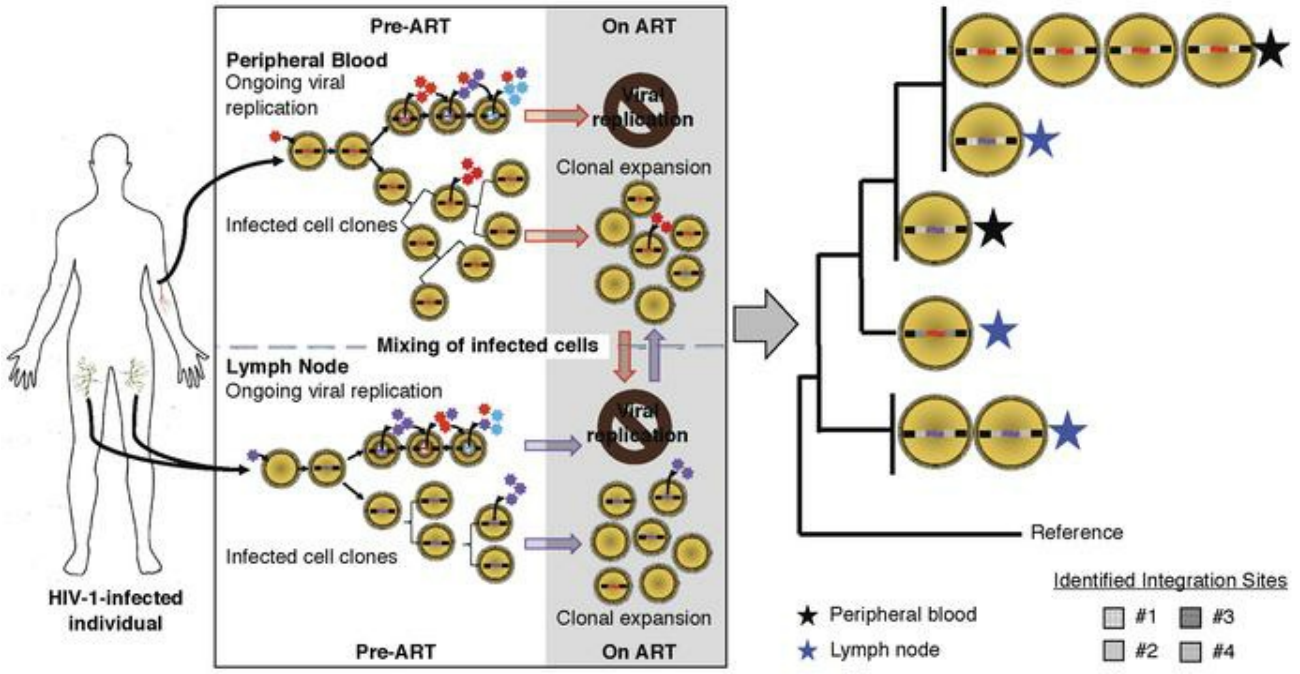
HIV-1 in lymph nodes is maintained by cellular proliferation during antiretroviral therapy

William R. McManus, … , John M. Coffin, Mary F. Kearney

J Clin Invest. 2019. https://doi.org/10.1172/JCI126714.

Research In-Press Preview AIDS/HIV

Graphical abstract



No evidence of ongoing viral replication in peripheral blood or lymph node

No compartmentalization between proviruses in peripheral blood and lymph node during long-term ART

Find the latest version:

http://jci.me/126714/pdf

HIV-1 in Lymph Nodes is Maintained by Cellular Proliferation During Antiretroviral Therapy

William R. McManus[1], Michael J. Bale[1], Jonathan Spindler[1], Ann Wiegand[1], Andrew Musick[1], Sean C. Patro[1], Michele D. Sobolewski[5], Victoria K. Musick[1], Elizabeth M. Anderson[1], Joshua C. Cyktor[5], Elias K. Halvas[5], Wei Shao[2], Daria Wells[2], Xiaolin Wu[2], Brandon F. Keele[2], Jeffrey M. Milush[3], Rebecca Hoh[3], John W. Mellors[5], Stephen H. Hughes[1], Steven G. Deeks[3], John M Coffin[4], Mary F. Kearney[1*]

[1]HIV Dynamics and Replication Program, CCR, National Cancer Institute, Frederick, MD, US
[2]Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD, US
[3]Department of Medicine, University of California, San Francisco, San Francisco, CA, US
[4]Department of Molecular Biology and Microbiology, Tufts University, Boston, MA, US
[5]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, US

*Corresponding Author:
Mary Kearney, PhD
HIV Replication and Dynamics Program
National Cancer Institute at Frederick
1050 Boyles Street, Building 535, Room 108D
Frederick, Maryland, US 21702
301 846 6796
kearneym@mail.nih.gov

Conflict of interest statement

JWM is a consult to and has received grants from Gilead Sciences and owns shares of Co-crystal Pharmaceuticals, Inc. The remaining authors have declared that no conflict of interest exist.

Abstract

To investigate the possibility that HIV-1 replication in lymph nodes sustains the reservoir during antiretroviral therapy (ART), we looked for evidence of viral replication in 5 donors after up to 13 years of viral suppression. We characterized proviral populations in lymph nodes and peripheral blood before and during ART, evaluated the levels of viral RNA expression in single lymph node and blood cells, and characterized the proviral integration sites in paired lymph node and blood samples. Proviruses with identical sequences, identical integration sites, and similar levels of RNA expression were found in lymph nodes and blood samples collected during ART, and no single sequence with significant divergence from the pretherapy population was detected in either blood or lymph nodes. These findings show that all detectable persistent HIV-1 infection is consistent with maintenance in lymph nodes by clonal proliferation of cells infected before ART and not by ongoing viral replication during ART.

Introduction

Antiretroviral therapy (ART) effectively prevents HIV-1 disease progression; however, low-level viremia persists during ART and typically rebounds to pre-therapy levels when treatment is discontinued. Rebounding virus is not genetically divergent from pre-ART HIV-1 variants (1-3), suggesting a stable viral reservoir during ART that is likely established prior to initiating treatment. Understanding the mechanisms that maintain the HIV-1 reservoir on ART is vital to developing strategies to eradicate the infection and/or to prevent viral rebound without ART.

Although there is ongoing debate (4, 5), evidence suggests that ART effectively blocks HIV-1 replication in the peripheral blood of both children and adults whether initiated in acute (6-8) or in chronic infection (3, 9, 10). These results imply, and previous experimental evidence supports, that HIV-1 proviruses in peripheral blood either reside in cells infected prior to ART or in their clonal descendants (8, 11-14). Such a model suggests, assuming that there is free exchange of infected cells between the lymph nodes and blood (15, 16), that there would be a lack of viral evolution in lymph nodes during ART and that the proviral population in the lymph nodes and peripheral blood would be highly similar. However, studies have claimed to be able to detect viral evolution in tissues collected during ART (16-18), suggesting that HIV-1 replication may not be fully inhibited in all anatomical sites. Furthermore, questions have been raised about the ability of ART to effectively penetrate lymph node follicles and prevent continuing rounds of viral replication (16, 19, 20), raising the possibility that HIV-1 may persist during ART by exploiting putative “sanctuary” sites in lymph nodes (19, 21, 22). The hypothesis that the HIV-1 reservoir is maintained during ART by persistent viral replication in lymph nodes implies that there is only minimal trafficking of infected cells between lymph nodes and peripheral blood. If not, then proviruses in infected cells in the blood would show evidence of evolution due to accumulating mutations, especially after many years of viral suppression on ART.

To determine if the mechanisms that maintain HIV-1 proviruses in lymph nodes differ from those in peripheral blood, and whether there is ongoing exchange of infected cells between these compartments, we compared the proviral populations in lymph nodes and peripheral blood from five participants whose levels of viremia were suppressed on ART (<40 copies/ml) for 1.8 to 12.9 years. We addressed the following questions related to the mechanisms that maintain the HIV-1 reservoir in lymph nodes during ART: 1) During long-term ART, are proviral populations in lymph nodes divergent from pre-ART proviral populations in lymph nodes and peripheral blood? 2) Are HIV-1 proviral populations divergent between paired lymph node and peripheral blood samples collected after long-term ART? 3) Do proviral populations in lymph nodes change over time in a manner characteristic of ongoing HIV-1 replication during ART? 4) Are the clonal populations of infected cells that persist on ART different in lymph nodes compared to peripheral blood? 5) Do more infected cells in lymph nodes contain HIV-1 RNA compared to those in peripheral blood? 6) Are proviruses in lymph nodes expressed at higher levels than those in peripheral blood? 7) Are the replication-competent proviruses in lymph nodes that are induced in viral outgrowth assays divergent from pre-ART viral populations? 8) Is there a higher number of CD4+ T cells carrying infectious, inducible proviruses in lymph nodes compared to peripheral blood during ART?

The results we obtained support the conclusion that HIV-1 infected cells persisting on long-term ART are well mixed between the peripheral blood and lymph nodes and that the HIV-1 reservoir is sustained in the lymph nodes by proliferation of cells that were infected prior to ART and not by ongoing cycles of viral replication during ART.

Results

Donor characteristics

Samples from five HIV-1 infected participants from the San Francisco SCOPE cohort (NCT00187512) (23) (Table S1) were investigated for evidence of ongoing HIV-1 replication, as assessed by evolution of sections of the viral genome in peripheral blood and lymph nodes during ART. In these donors, viremia was well-suppressed on ART (<40 HIV-1 RNA copies/ml plasma) for 1.8 to 12.9 years. Two participants, 3720 and 2661, achieved viral suppression in early infection (<1 year and 0.3 years after the approximate date of transmission). The low pretreatment genetic diversity of the HIV-1 populations in these donors provided a sensitive probe for the detection of viral divergence in the peripheral blood or lymph nodes for periods of 1.8 years and 12.9 years, respectively. Donor 2661 had one brief ART interruption after 4.7 years, was quickly re-suppressed, and subsequently maintained continuous suppression of viremia on ART for another 12.9 years. The other three participants initiated ART in chronic infection (2.0 to >22 years after the approximate date of transmission).

All donors had HIV-1 subtype B infection except donor 3720 who was infected by subtype C. With one exception, no participant had detectable mutations conferring resistance to current therapy in any samples studied. Donor 2669 had a history of mono- and/or dual-therapy, and, consequently, acquired reverse transcriptase inhibitor drug resistance mutations prior to achieving continuous viral suppression for 5.5 years following the addition of dolutegravir to his treatment regimen. Single-genome sequencing (SGS) of the integrase coding region in peripheral blood mononuclear cells (PBMC) and lymph node mononuclear cells (LNMC) did not detect mutations that would contribute to dolutegravir resistance (92 single-genome DNA sequences obtained from PBMC and 115 from LNMC).

To compare the evolution of HIV-1 between the different locations, pre-ART plasma and/or PBMC and/or LNMC were obtained from four of the participants. Two longitudinal on-ART LNMC samples were obtained about 1 year apart from two of the participants (2669 and 1079), including

paired samples from contralateral inguinal lymph nodes, enabling comparison of the proviral populations in the two nodes. After completing this study, one donor interrupted ART, and a plasma sample was obtained 14 days after discontinuation, allowing the rebounding plasma viral RNA population to be compared to pre-ART RNA sequence variants in plasma, to the pre-ART DNA and cell-associated RNA variants in peripheral blood, and to the on-ART DNA and cell-associated RNA variants in peripheral blood and lymph nodes. A detailed timeline of participant sampling is shown in Table S2.

No evidence for HIV-1 evolution during ART in peripheral blood

Previous studies have demonstrated the absence of detectable HIV-1 evolution in the peripheral blood during ART (3, 6, 9, 10). Because HIV-1 replication is error prone, this result shows that there is, at most, only very limited HIV-1 replication in blood. To ensure that ART fully suppressed viral replication in the peripheral blood of this set of donors, we performed SGS of the P6-PR-RT region derived from virus in the plasma or viral DNA in PBMC collected prior to ART and after 1.8 to 12.9 years of continuous viral suppression (Figure 1). We also performed SGS on the full-length *env* gene in proviruses from pre- and on-ART PBMC samples from donors 1683 and 1079 (Figure S1). Because evolution would be a consequence of ongoing viral replication, we looked for evidence of viral evolution on ART. Evolution of the proviral sequences would cause: 1) increasing genetic diversity measured by average pairwise distance (APD), 2) increasing genetic divergence of individual sequences from the starting population over time, as shown by a statistically significant ($p<0.001$) negative test for panmixia comparing the two time points (24), and 3) emergence of new viral variants, observable by increasing average and individual root-to-tip distances over time on maximum-likelihood phylogenetic trees (Figure 1, S1).

There was no evidence of HIV-1 evolution characteristic of ongoing cycles of viral replication detected (in either P6-PR-RT or *env*) in peripheral blood from any of the four individuals we studied by any of these analytical tools. Two donors (Figure 1A, B) were treated in early

infection and the HIV-1 populations did not shift from the apparent pre-ART founder viruses over 1.8 and 12.9 years of continuous treatment (probability of panmixia = 0.1 and 0.4, respectively). HIV-1 sequences with APOBEC-mediated G to A hypermutation, which are commonly seen in individuals on ART (25), as well as any other sequences containing stop codons, were excluded from these and all subsequent analyses. In donor 1683 (Figures 1C and S1A), the diversity in the proviral population decreased during ART in both P6-PR-RT and *env* (0.5% to 0.1%, $p<0.0001$ and 1.3% to 0.6%, $p<0.0001$). In this donor, the HIV-1 population became dominated by a single clonal sequence that was already present in the pre-ART population 5.4 years prior and may have been directly derived from the original founder virus. The increased proportion of this clonal variant also resulted in divergent (non-panmictic) proviral populations prior to and during ART in P6-PR-RT ($p=5x10^{-6}$), a difference that disappeared when the identical sequences were collapsed to single variants ($p=0.7$), as was done for all the panmixia analyses shown in the figures. This result implies that the apparent shift in the population during ART was due to clonal proliferation of infected cells, rather than emergence of new variants from ongoing replication. The root-to-tip analysis also showed a potential loss of P6-PR-RT sequences on longer branches from the pre-ART population (slope = $-2.0x10^{-5}$). Similar results were obtained from analyses of the same two genomic regions in peripheral blood from donor 1079 (Figure 1D and S1B), treated in chronic infection, despite the absence of prominent clonal variants in these samples after 11.4 years on ART.

As in previous studies (3, 6, 9, 10), comparative analyses of single-genome sequences obtained before therapy and after up to 13 years of suppressive ART provided no evidence for ongoing replication of HIV-1 in peripheral blood. On the contrary, all changes in the genetic structure of proviral DNA populations can be attributed to clonal proliferation of cells containing proviruses laid down prior to treatment initiation or the loss of variants during ART.

No evidence for HIV-1 evolution during ART in lymph nodes

It has been suggested that the drugs used in ART do not optimally penetrate lymph nodes (19, 22), permitting persistent HIV-1 replication in lymph nodes (16, 20). If significant levels of viral replication occur in the lymph nodes during ART, then the newly emergent viral variants produced by the accumulation of reverse transcriptase errors would have to be sequestered in the lymph nodes, because they are not detected in the peripheral blood (as shown above). This argument predicts significant compartmentalization of HIV-1 DNA between the infected mononuclear cells in the lymph nodes and those in the peripheral blood. To address the possibility that HIV-1 infected cells are compartmentalized between the lymph nodes and peripheral blood, and that viral replication continues during ART at levels that sustain the HIV-1 reservoir in the lymph nodes as previously claimed (16), we compared proviral sequences of P6-PR-RT (Figure 2) and *env* (Figure S2) from the lymph nodes and peripheral blood of the donors in this study.

The results of this analysis yielded no evidence of compartmentalization in any of the donors when proviral sequences from the paired peripheral blood and lymph node samples were compared, in aggregate or individually, using the same three analytical tools as above (Figure 2, S2). These results (in detail below) indicate that infected cells are well-mixed between the two compartments, and imply the absence of significant viral replication in lymph nodes, just as in peripheral blood, during suppressive ART. Neighbor joining phylogenetic analyses of P6-PR-RT sequences from the two donors who were treated in early infection (Figures 2A-B) showed, in each case, a prominent group of identical sequences in the lymph nodes that was identical to the presumed founder viruses (the most common variants observed in the peripheral blood prior to ART), suggesting that no detectable sequence divergence occurred in the lymph nodes over 1.8 or 12.9 years of viral suppression on ART in these two individuals (there were not enough cells available from these donors to sequence *env*). After sampling over half of the infected cells in an entire lymph node from donor 2661, we found only a few P6-PR-RT variants to be 1-2 nucleotides different from the sequence of the presumed founder virus and these had not diverged from the plasma virus in the pre-ART sample (APD: $p=0.08$, panmixia: $p=0.2$). Though the positive branch

length correlation coefficient indicated a significant difference in branch lengths between lymph node and peripheral blood (r=0.26; p=0.0005), when identical sequences were converted into single variants this difference disappeared (r=-0.065; p=0.6), suggesting that the difference was due to the fact that an identical variant was sampled at a disproportionately higher level in the peripheral blood than in lymph nodes or that an expanded clone carrying this variant is not equal in size across the two locations. As in the blood, the only noticeably divergent sequences in the lymph nodes were G to A hypermutants.

A lack of compartmentalization between peripheral blood and lymph nodes was also observed in the three individuals treated in chronic infection (Figures 2C-E, S2A-C). Sequences obtained from paired peripheral blood and lymph nodes were well-mixed phylogenetically and by the test for panmixia, had similar measures of diversity, and had branch length correlations not significantly different from zero in both P6-PR-RT (Figure 2C-E) and *env* (Figure S2A-C), with the exception of donor 1683. There was a slight increase in the diversity of the proviral P6-PR-RT sequences in the lymph node from this donor (from 0.1 to 0.3%, p=0.04); however, as in donor 2661, this difference can be attributed to the disproportionate representation, between the two locations, of a major clonal variant, a difference that likely arises from the physiology of cells in those particular locations (such as expression of homing signals or responsiveness to antigen (14)) rather than a property of the provirus itself. When identical sequences were condensed into single variants, the diversity of the populations in the peripheral blood and lymph nodes was not statistically different (p=0.7). Differences in the number of rakes of identical sequences were observed between P6-PR-RT and *env*, particularly in PID 1683 (Table 2). Such differences likely reflect the more rapid accumulation of diversity in *env* compared to *pol* and, possibly, the different selection pressures on *env* vs. *pol* including the humoral immune response and/or cell tropism. Identical sequences common to both peripheral blood and lymph nodes (indicated with black arrows in the figures) were found in the donors who initiated ART in chronic infection; these identical sequences likely resulted from clonal expansion of cells infected prior to the initiation of

ART. If so, these data imply trafficking of clonally expanded cells between the two compartments. The presence of indistinguishable HIV-1 populations in paired peripheral blood and lymph node samples collected during ART suggests, with our level of sampling (see discussion), that ongoing HIV-1 replication in lymph nodes is likely insufficient to sustain the HIV-1 reservoir during treatment.

In addition to seeking evidence for compartmentalization between lymph nodes and peripheral blood, we also investigated the possibility of compartmentalization between two separate lymph nodes from the same donor at the same time point (Figure S3). LNMC were collected from contralateral inguinal lymph nodes of two donors and proviral P6-PR-RT sequences were obtained. In both donors, we found the HIV-1 populations in the paired lymph nodes to be well-mixed, showing no evidence for viral compartmentalization between nodes by any metric (for donors 1079 and 2669, respectively: APD: p=0.8 and 0.3, panmixia: p=0.3 and 0.7, branch length correlation: p=0.05 and 0.7). These data imply that infected cells traffic between peripheral blood and lymph nodes and provide additional evidence that ART effectively blocks HIV-1 replication in lymph nodes as well as peripheral blood to levels that are not detectable with our level of sampling by SGS.

To further address the question of ongoing viral replication in lymph nodes during ART, we pooled the proviral sequences obtained from the paired lymph node samples in each donor and compared those sequences to proviral sequences obtained from a lymph node one year prior (during ART) or from a lymph node collected prior to ART initiation (Figure 3). Using the analytical tools described above, we observed no change in the diversity or divergence of the HIV-1 proviral populations in the longitudinal lymph node samples (for donors 3720, 1079, and 2669, respectively: APD comparison: p=0.8, 0.9, and 0.9, panmixia: p=0.6, 0.8, and 0.9, and branch length correlation: p=0.9, 0.9, and 0.9). The lack of any detectable change in the proviral populations in longitudinal lymph nodes during years of therapy again shows that ART is effective at blocking viral replication in the lymph nodes.

HIV-1 Integration site comparison of clonal populations in lymph nodes and peripheral blood

The sites at which HIV-1 proviruses are integrated can be used to monitor the clonal expansion of infected cells (11). Because HIV-1 integrates almost randomly into many sites in the human genome, finding multiple cells with exactly the same sites of integration is very strong evidence that these cells are descendants from a single infected cell. By this principle, integration site analysis can be used to monitor the trafficking and compartmentalization of populations of HIV-1 infected cells. Thus, when infected cells at two different locations have identical integration sites, they must be members of a clone derived from the same infected cell, and cells from that clone must have trafficked between the two locations. We used the integration sites assay to determine whether the same expanded clonal populations were present in lymph node and peripheral blood (Table 1). From each donor, we obtained 143-247 integration sites from the peripheral blood samples and 63-579 integration sites from the lymph nodes [full set of integration sites deposited in the Retroviral Integration Sites Database (RID), accessible by Pubmed ID at (https://rid.ncifcrf.gov/)]. Clones were defined as proviruses with identical integration sites that were detected at least once in each of two different locations or samples regardless of ART status, in two or more different DNA fragments in the same location or sample during ART, or three or more times prior to ART (26). Although the presence of two identical integration sites is indicative of cell division, we define clones in pre-therapy conservatively as requiring three detection events because dividing cells could die before establishing an expanded cell clone. In the three ART treated donors in which the numbers of infected cells were sufficient for integration site analyses (1683, 1079, and 2669), common clones were detected between lymph nodes and peripheral blood. Of note, in the combined dataset, we found evidence for selection for integrations in the *BACH2* ($p=3.0\times10^{-7}$), *MKL1* ($p=0.010$), and *STAT5B* genes ($p=3.5\times10^{-4}$) as previously reported (11, 13). Overall, in these three donors, the clonal populations were not significantly different

between the lymph nodes and peripheral blood (p=0.1), confirming the absence of detectable compartmentalization of infected cells between those locations (Table 2).

Expression of HIV-1 proviruses in peripheral blood and lymph nodes prior to and on ART

Our analysis of proviral populations showed no evidence of ongoing HIV-1 replication in, or compartmentalization between, peripheral blood and lymph nodes. In a further study, we determined the fraction of infected cells with unspliced HIV-1 RNA (monitoring P6-PR-RT) and the levels of proviral expression in single cells in each location. Using our cell-associated RNA and DNA single-genome sequencing (CARD-SGS) assay, which can detect a single HIV-1 P6-PR-RT RNA sequence in a single infected cell (27), we compared the level of unspliced viral RNA in single cells obtained from paired peripheral blood and lymph node samples (Table 3, Table S3) (Figure 4, 5, S4). While our previous work showed that there are single cells with high levels of HIV-1 RNA in the peripheral blood of untreated individuals, cells with similarly high levels of viral RNA were not detected in the peripheral blood of patients on ART (27). The observation of cells containing high levels of viral RNA (which may be on the viral replication pathway) in lymph nodes during ART would provide support for the proposal that there are significant levels of ongoing replication in the lymph nodes. Although cells expressing high levels of HIV-1 RNA can be seen in in situ hybridization assays on lymph nodes of HIV-infected individuals on ART (20, 28, 29), they are much more rare than in untreated individuals, and may well reflect activation of HIV-1 expression in descendants of cells infected prior to ART initiation. The low-level persistent viremia seen in most suppressed patients implies that highly-expressing cells must exist, albeit at numbers averaging about 10,000-fold lower than before treatment initiation (27).

We found no overall difference in the fraction of infected cells that contained viral RNA in lymph nodes compared to peripheral blood during ART (median 13% in LNMC vs. median 8% in PBMC) (p=0.13) or in the levels of proviral expression within single cells between the lymph nodes and peripheral blood (median = 1 RNA molecule/expressing cell in both). In donor 3720, we

detected a few cells in the peripheral blood and lymph node with high levels of viral RNA (containing >20 HIV-1 RNA molecules) prior to ART, while no such cells were observed in a sample taken after 1.8 years on ART (Figure 5A, cells with high viral RNA levels indicated with blue arrows; pre-ART data from lymph nodes, including 3 cells with high HIV-1 RNA levels, not shown). Similarly, in donor 1683, one cell with a high level of viral RNA was identified in the peripheral blood prior to ART, and none was found after 5.4 years on ART in either peripheral blood or lymph nodes (Figure 5B, indicated with blue arrow). In donor 1079, no cells with high levels of viral RNA were detected at either time point (Figure 5C, 353 HIV-1 infected cells assayed for HIV-1 RNA from pre-ART and 1315 on ART). Although a pre-ART sample was not available for donor 2669, we did detect one LMNC during ART with 28 molecules of HIV-1 RNA (Figure S4C). These data suggest that cells that have high levels of viral RNA are rare and, as expected for productive infection, may be preferentially eliminated. It is also likely that latently-infected cells, or cells with very low levels of HIV-1 expression, are not eliminated and, therefore, accumulate over the duration of infection, possibly explaining why high-expressing cells were detected before ART initiation in donors 3720 and 1683 (who were still rather early in infection) and not in 1079 (who had been infected for more than 4 years at the time of ART initiation), although a larger study would be required to investigate this possibility.

As reported previously (27), we also found RNAs with identical P6-PR-RT sequences in different single cells, suggesting that clonally proliferating populations of infected cells include ones with proviruses that are actively expressing low levels of HIV-1 RNA. While it is possible that such cells are the source of viral rebound when ART is interrupted, it is more likely that there are always very rare highly expressing cells, as can be seen by in situ hybridization (20, 28, 29) that give rise to persistent and rebound viremia and, in some cases, are the source of outgrowth in viral outgrowth assays. In donor 3720, the viral outgrowth assay (30) identified multiple sequences obtained from pre-ART as replication competent (Figure 1A, 2A, 3A indicated with red arrows). Two of the viral sequences obtained in viral outgrowth assays matched RNA sequences obtained

from multiple PBMC prior to ART and from multiple PBMC and LNMC obtained after 1.8 years on ART (Figure 5A, S4A, indicated with red arrows). In donor 1079, a treatment interruption revealed two major rebound variants in the consequent viral rebound, one of which matched a set of six lymph node-derived cell-associated RNA sequences obtained from a single cell after 11.4 years on-ART (Figure 4B, indicated with yellow arrow). This observation suggests that rebound viremia resulted from a clone of infected cells, some of which were actively transcribing HIV-1 RNA prior to ART interruption, as previously reported (2). The other rebounding variant matched a single proviral sequence obtained from the peripheral blood during ART (data not shown but sequence submitted to GenBank, accession #MK147615). Thus, persistent, rebound, and rescued virus can have the same genetic properties as the proviral DNA described above, showing no evidence of ongoing replication or evolution, but strong evidence for clonal expansion of the cells responsible for it.

In estimating the fraction of infected cells that contain HIV-1 P6-PR-RT RNA, our measurements of the number of HIV-1 infected cells were based on a particular region of the genome, in the *pol* gene (31) (Table 3). To ensure that the region of the viral genome we analyzed did not affect our conclusions, we also used primers and probes in the LTR (R-U5) and in the *gag* gene (Table S3). Though quantification by these methods resulted in varying counts of infected cells (likely due to some proviruses having large internal deletions or variants with mismatches to the primers or probes), our estimates for the fractions of infected cells with HIV-1 P6-PR-RT RNA were not significantly different between the peripheral blood and lymph nodes, regardless of the genomic region used to detect proviral sequences ($p=0.4$ for LTR and $p=0.7$ for *gag*). Because there are more CD4+ T cells in lymph nodes of some donors, and more infected cells in the lymph nodes than in the peripheral blood, the overall number of HIV-1 RNA molecules we detected could be higher in lymph nodes than in peripheral blood, as previously reported (20, 28). However, when the levels of HIV-1 RNA in single infected cells were compared, we found no difference between the cells from peripheral blood and lymph nodes during long-term ART. These findings

demonstrate that lymph nodes cells containing high levels of HIV-1 RNA collected from HIV-infected individuals on ART, such as those detected here and previously using in situ hybridization assays (27-29), more likely reflect activation of HIV-1 expression in descendants of cells infected prior to ART initiation than from cells infected from ongoing cycles of viral replication on ART.

No single proviral sequences on ART deviate from those in pre-ART

Our results clearly show that there is no detectable evolution of the bulk population in either peripheral blood or lymph nodes for almost 13 years of suppressive ART. However, it remains possible that a fraction of the proviral DNA population may have arisen by low-level ongoing replication and that these sequences may not have been visible in analyses of the population as a whole. Such variants would only be visible in the populations when analyzed individually, as distinct outliers in the overall distribution of a diversity metric. To test for the presence of such variants in the combined patient data, we normalized the root-to-tip distances of all HIV-1 RNA or DNA sequences from pre-ART in each patient, when available, and divided the results into bins, yielding a histogram centered around 0, as shown in Figure 6 (the red bar graph). We then did the same analysis with the P6-PR-RT proviral sequences from the on-ART peripheral blood and lymph node samples, normalizing individual distances to the respective donor's pre-ART data, as shown in Figure 6 (grey and blue graphs, respectively). Thus normalized, all patient data could be aggregated, allowing visualization of evolution of the peripheral blood and lymph node populations overall, as well as sensitive detection of small subpopulations that might be evolving separately from the bulk population. Analyses were performed both on total sequences (left panels) and with groups of identical sequences collapsed to remove the effects of clonal expansion (right panels).

Consistent with our previous analyses, the pre-ART and on-ART datasets yielded very similar histograms, despite the samples being drawn from different anatomical locations and separated by very long times. In the plots, any single sequence with a genetic distance of more

than two standard deviations from the mean of the pre-ART sequence distribution would be revealed as a significant outlier. Since we examined a total of about 125 sequences in each set, we expected about 5% of the 122 PBMC or 129 LNMC sequences to be between 2 and 3 standard deviations from the mean by chance, and none to be greater than 3. As the figure shows, the number of such outlier sequences is actually less than the 6 sequences expected (0/122 total sequences in the on-ART PBMC and 1/129 sequences in the on-ART LNMC) and very similar to the frequency of such outliers in the pre-ART data. Given that HIV-1 genomes accumulate mutations at about 1% per year in untreated patients (32) and one standard deviation in this analysis corresponds to a genetic distance of between 0.08 and 0.9% from the root in these donors, 3 standard deviations represents a very low bar based on the time between ART initiation and the on-ART sampling, yet not a single provirus in 251 studied in the two tissues was able to cross it. Thus, if there is ongoing replication, it can only involve, at most, 0.4% (upper bound of 95% CI = 1.2%) of the virus population in either blood or lymph nodes in ART-suppressed patients.

Discussion

The question of whether ongoing viral replication sustains the HIV-1 reservoir during ART has important implications for the design of curative interventions. Some have concluded that viral replication in lymph nodes continues at levels that maintain the HIV-1 reservoir (16), and others have claimed that viral replication may persist during ART, but in only a small subset of cells (20, 33). If HIV-1 replication persists on ART to the extent necessary to maintain the viral reservoir, then the accumulation of new proviral mutations would be apparent after long-term ART. The acquisition of new mutations would be especially evident in patients who initiated treatment in early infection with a homogeneous virus population, such as donor 2661. In this donor, the number of HIV-1 infected cells that persisted on ART was small compared to the individuals who initiated ART in chronic infection, consistent with previous reports (34-37). As a result, all 150 million LNMC obtained from an entire lymph node were investigated for HIV-1 DNA or RNA. We found that <600 infected cells persisted in the lymph node from this individual after almost 13 years on ART and <13% had unspliced HIV-1 RNA. Extrapolating this finding to the ~500-700 lymph nodes in the body, we can estimate that about 300,000-420,000 HIV-1 infected LNMC persisted in this individual and about 39,000–55,000 had unspliced HIV-1 RNA. In contrast, the donors who initiated ART in chronic infection and had been sustained on ART for a shorter duration (about 1-5 years), according to our calculations, had as many as 175 million infected LMNC in the entire body and as many as 22.8 million with unspliced HIV-1 RNA, 456,000 of which may be “high-expressers”. In total, we examined 251 lymph node and peripheral blood proviral-derived sequences from the donors who had pre-ART sampling and we failed to detect a single provirus with a significant number of new mutations in either location in any of the donors. Thus, if there is a fraction of the proviral population in either lymph node or peripheral blood representing virus generated by ongoing replication, it must be very small. With our level of sampling, we can report with 95% confidence that, if there is a replicating population, it is <1.2% of the total HIV-1 infected cells that persist on ART. Additionally, as viral RNA+ cells do not necessarily indicate

replication competency, <1.2% is a highly conservative estimate of the upper bound of a potential replicating pool (25, 38).

By contrast, the clonal proliferation of infected cells that contain replication-competent proviruses can generate large pools of infected cells, some of which persist during long-term ART (14, 39-41). Our integration site assay data shows that clonal populations of infected cells that are detectable after long-term ART in the peripheral blood and lymph nodes are not significantly different and provides support for the conclusion that cell proliferation is likely to be the most important mechanism that maintains the persistent viral DNA in blood and lymph nodes. These data, taken together with the finding that the HIV-1 populations in peripheral blood and lymph nodes are not genetically divergent and that infected cells in those locations contain similar levels of viral RNA, show that this process is not different in these compartments. Importantly, these data also show that infected cells are trafficked between the blood and the lymph nodes, as well as between lymph nodes. This lack of compartmentalization further implies that ongoing cycles of viral replication during ART do not persist in lymph nodes at a level that sustains the HIV-1 reservoir. Our findings support the conclusion that the viral DNA reservoir is maintained largely or entirely by clonal proliferation of cells that were infected prior to initiating ART and that this mechanism is important in both the peripheral blood and the lymph nodes. In sum, our results refute the idea that development of new antiretroviral drugs (with improved targeting to lymph nodes, for example) will be of value towards achieving a cure of HIV-1 infection. The very low frequency of virus producing cells detectable by in situ PCR assays (20, 28, 29), may be important biologically for other reasons not addressed in this study, but our results imply that they are more likely to be due to expression of rare proviruses in descendants of cells infected prior to therapy than from ongoing viral replication. Likewise, the low number of high-expressing cells, though likely to be the source of persistent viremia on ART and rebound viremia after stopping ART, is unlikely to result in ongoing cycles of viral replication in the presence of effective therapy. These results make it very clear that survival and expansion of cells infected prior to therapy constitute

451 the major, if not the only, barrier to a cure of HIV-1 infection. Future strategies for curative
452 interventions must focus on methods to target the populations of clonally expanded HIV-1 infected
453 cells harboring intact HIV-1 proviruses.
454

Methods

Participant cohort and sample collection

Samples from the five HIV-1 infected participants were obtained from the San Francisco SCOPE cohort (NCT00187512) (23) (Table S1). PBMC were separated with Ficoll and resuspended in FBS with 10% DMSO. To disaggregate LNMC from lymph node biopsies, excess fat was trimmed and the node was minced and strained on 70 and 40 micron nylon cell strainers. Samples were shipped on dry ice and stored in liquid nitrogen until processed.

Sequence data availability

All sequence data are available on GenBank at accession numbers MK145079-MK148700.

HIV-1 quantification

HIV-1 DNA levels were determined using the integrase cell-associated DNA assay (31) or using a droplet digital PCR (ddPCR) assay (42). For ddPCR, HIV-1 DNA copies per million cells were measured in triplicate with a duplexed ddPCR assay measuring the amount of DNA corresponding to HIV-1 *gag (43), pol,* and the RU5 region of the HIV-1 LTR. The amount of cellular DNA was measured by CCR5 quantitation. The RU5 region of the HIV-1 LTR was amplified using the following primers: RU5-F- 5'-CTTAAGCCTCAATAAAGCTTGCC-3', RU5-R- 5'-GGATCTCTAGTTACCAGAGTC-3', and RU5-Probe (Hex Zen)-5'-AGTAGTGTGTGCCCGTCTG-3'. Prior to amplification, DNA was extracted as in (31) and randomly sheared with a Branson ultrasonic cup horn sonifier (Emerson) at 60% amplification in pulse mode for 5 seconds. This step was repeated 3 times. Samples were then heated to 100^{o}C for 15 minutes and snap cooled on ice. A 20µl PCR mastermix was made which gave a final

concentration of 1x ddPCR Supermix for Probes (BioRad, USA), 750nM forward and reverse primers, 250µM probe and 5µl DNA template. End-point PCR was performed with the following conditions: 95^{o}C for 10 minutes, then 40 cycles of 94^{o}C for 30 seconds, 55^{o}C for 1 minute with a ramp rate of 2^{o}C/second, followed by a final 98^{o}C for 10 minutes. After the reactions were completed the samples were held at 12^{o}C. Reactions were then read on the QX200 Droplet Reader (Bio-Rad, USA) and were analyzed using the Quantasoft software version 1.7.4 (Bio-Rad, USA) with a user defined threshold. Averaged triplicate HIV DNA measurements were normalized to one million PBMCs using concurrent triplicate CCR5 measurements.

Cell-associated RNA and DNA single-genome sequencing assay (CARD-SGS)

CARD-SGS was performed on the P6-PR-RT region of HIV-1 as previously described (27). In brief, CARD-SGS is performed by diluting cells to near the endpoint for those with HIV-1 unspliced RNA and performing single-genome sequencing on reverse transcribed RNA. Using single ACH2 cells, we previously showed that our method can detect a single HIV-1 P6-PR-RT RNA molecule in a single cell (26). Because the reverse transcription step is known to introduce errors at a rate of about 10^{-4} per sequenced nucleotide of viral cDNA sequences (44), single RT-PCR variants that differed by a single nucleotide from a group of 5 or more identical sequences within the same aliquot were counted with the majority. HIV-1 proviral SGS was performed on full-length *env* as previously described (45).

SGS of the full-length integrase coding region was carried out on PBMC cell-associated DNA using the same PCR parameters as those described for P6-PR-RT (46) but using different PCR primers: IntF1- 5'-CATCTAGCTTTGCAGGATTCG-3' and IntR1- 5'-CTGACCCAAATGCCAGTCTC-3' and for nested PCR IntF2- 5'-GGAAAAGGTCTACCTGGCATG-3' and IntR2- 5'-TCTCCTGTATGCAGACCCCA-3'. Sequences

were aligned using ClustalW and all sequences that would clearly render the virus defective for replication (stop codons within the coding region) were omitted from downstream analyses. Population genetic diversity was calculated as average pairwise p-distance (APD) using MEGA (www.megasoftware.net). Neighbor joining trees were constructed using MEGA and rooted on consensus B or C as appropriate.

Integration sites assay (ISA)

ISA was performed as previous described (11) using patient-specific primers to the 5' and 3' LTRs. The full set of integration sites obtained were submitted to the Retroviral Integration Sites Database (https://rid.ncifcrf.gov/) and can be accessed using the Pubmed ID for this paper. The primer sequences are as follows:

3720 PCR 1
3'LTR: TGTGGACTCTGGTAACTAGAGATCCCTC
5'LTR: TCAGGGAAGTAGCCTTGTGTGTGGT

3720 nested PCR
3'LTR: CCCTTTGTGGTAAGTGTGGAAAATC
5'LTR: TCTTGGCTCTTCTTGGAGTAAACTA

1683 PCR 1
3'LTR: TGTGACTCTGGTACTAGAGATCCCTC
5'LTR: TCAGGGAAGTAGCCTTGTGTGTGGT

1683 nested PCR
3'LTR: CCCTTTTAGTCAGTGTGAAAATC
5'LTR: TCTTGCCTTTGCTGGGAGTAAATTA

1079 PCR 1
3'LTR: TGTGACTCTGGTACTAGAGATCCCTC
5'LTR: TCAGGGAAGTAGCCTTGTGTGTGGT

1079 nested PCR
5'LTR: CCTATTTAGTCAGTGTGGAAAATC
3'LTR: TAAAAAGTGGCTAAGATCCAGAGC

2669 PCR 1
3'LTR: TGTGACTCTGGTAACTAGAGATCCCTC
5'LTR: TCAGGGAAGTAGCCTTGTGTGTGGT

2669 nested PCR
5'LTR: CCCTTTTAGTCAGTGTGGAAAATC
3'LTR: CTTGTCTTTTCTGGGAGTGAATTA

Signed Relative Deviation Analysis of Single Proviruses

Root-to-tip distances from patient ML trees containing all pre-ART sequences, pre-ART sequences with identical sequences collapsed to a single variant, all on-ART proviral sequences from PBMC and LNMC, and on ART proviral sequences from PBMC and LNMC with identical sequences collapsed to a single variant were generated by phyML as described above. Normalization was done by taking the average and standard deviation of root-to-tip distances from the pre-ART sequence data. Signed relative standard deviations (sRSD) was then calculated as $\mathrm{sRSD_i} = \frac{\mathrm{RTT_i} - \overline{\mathrm{RTT}}_{\mathrm{pre-ART}}}{\sigma_{\mathrm{pre-ART}}}$ for each sequence *i* in each patient. The sRSD were then plotted as histograms, as shown in Figure 6. The upper limit of the confidence interval was determined by the "rule of three" for an unseen event (47) .

Statistics

Implementation of the Hudson test for geographic subdivision—a test for panmixia—was run using an in-house program and a significance cutoff of $<10^{-3}$ was applied to account for multiple comparisons (24, 48). Maximum likelihood (ML) trees were generated using the best fit model from the Smart model selection and estimated in PhyML (49, 50). Trees were rooted against consensus HIV-1 subtype B or C sequences as appropriate. Root-to-tip distances were plotted against the time of sample collection and evaluated via F-test on the resulting linear regression (units for root-to-tip are substitutions/year). Branch length correlation coefficients for compartmentalization were determined calculated as described in Critchlow *et al.* using an in-

house program (51). Differences in APD were tested as described previously (6). Other standard statistical analyses and summary statistics were done in R version 3.3.1. and are noted in figure legends. In-house programs for the test for panmixia (java), root to tip analysis (python3), and branch length tree correlation coefficient (python3) are available online at www.github.com/michaelbale.

Study approval

Samples from the five HIV-1 infected participants were obtained from the San Francisco SCOPE cohort (NCT00187512) (23) (Table S1) and studies were approved by the internal review board of the University of California San Francisco. Written informed consent was received from the participants prior to inclusion in the study.

Author contributions

WRM – performed SGS, analyzed data, wrote the paper
MJB – analyzed data, performed statistical analyses, wrote the paper
JS – performed SGS, analyzed data
AW – performed SGS
VM – performed SGS, analyzed data
AM – performed SGS, performed ddPCR, analyzed data
SCP – wrote the paper
MS – performed VOA
EMA – developed ddPCR assay
JC – performed iCARD assay, performed FACS analyses
EH – performed iCARD assays
WS – analyzed data
DW – performed integration sites assay (ISA)
XW – analyzed data
BFK – performed SGS
JM – processed patient samples; FACS analyses
RH – enrolled patients; analyzed data; wrote the paper
JWM – wrote the paper, contributed data
SHH – analyzed data, wrote the paper
SGD – conceived of idea, enrolled patients, wrote the paper
JMC – conceived of idea, analyzed data; wrote the paper
MFK – conceived of idea, analyzed data, wrote the paper

Acknowledgments

We thank Connie Kinna, Anne Arthur, Valerie Turnquist, and Sue Toms for administrative support, Valerie Boltz and Frank Maldarelli for useful discussions, and Adam Capoferri and Joe Meyer for help with graphics. We also thank the Viral Evolution Core in the AIDS and Cancer Virus Program; Christine Fennessey, Leslie Lipkey, Laura Newman, and Carolyn Reid for sequencing support.

Competing Financial Interests:

JWM is a consultant to Gilead Sciences, a shareholder of Co-Crystal Pharmaceuticals, Inc, and has received research support from Gilead Sciences and Janssen Pharmaceuticals. The remaining authors have no potential conflicts.

References

1. Joos B, Fischer M, Kuster H, Pillai SK, Wong JK, Böni J, et al. HIV rebounds from latently infected cells, rather than from continuing low-level replication. *Proceedings of the National Academy of Sciences of the United States of America.* 2008;105:16725-30.
2. Kearney MF, Wiegand A, Shao W, Coffin JM, Mellors JW, Lederman M, et al. Origin of Rebound Plasma HIV Includes Cells with Identical Proviruses That Are Transcriptionally Active before Stopping of Antiretroviral Therapy. *J Virol.* 2015;90(3):1369-76.
3. Kearney MF, Spindler J, Shao W, Yu S, Anderson EM, O'Shea A, et al. Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy. *PLoS Pathog.* 2014;10(3):e1004010.
4. Rosenbloom DIS, Hill AL, Laskey SB, and Siliciano RF. Re-evaluating evolution in the HIV reservoir. *Nature.* 2017;551(7681):E6-E9.
5. Kearney MF, Wiegand A, Shao W, McManus WR, Bale MJ, Luke B, et al. Ongoing HIV Replication During ART Reconsidered. *Open Forum Infect Dis.* 2017;4(3):ofx173.
6. Van Zyl GU, Katusiime MG, Wiegand A, McManus WR, Bale MJ, Halvas EK, et al. No evidence of HIV replication in children on antiretroviral therapy. *J Clin Invest.* 2017.
7. Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, et al. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci U S A.* 2013;110(51):E4987-96.
8. von Stockenstrom S, Odevall L, Lee E, Sinclair E, Bacchetti P, Killian M, et al. Longitudinal Genetic Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool During Effective HIV Therapy. *J Infect Dis.* 2015;212(4):596-607.
9. Vancoillie L, Hebberecht L, Dauwe K, Demecheleer E, Dinakis S, Vaneechoutte D, et al. Longitudinal sequencing of HIV-1 infected patients with low-level viremia for years while on ART shows no indications for genetic evolution of the virus. *Virology.* 2017;510:185-93.
10. Brodin J, Zanini F, Thebo L, Lanz C, Bratt G, Neher RA, et al. Establishment and stability of the latent HIV-1 DNA reservoir. *Elife.* 2016;5.
11. Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, et al. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science.* 2014;345(6193):179-83.
12. Bailey JR, Sedaghat AR, Kieffer T, Brennan T, Lee PK, Wind-Rotolo M, 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+ T cells. *J Virol.* 2006;80(13):6441-57.
13. Wagner TA, McLaughlin S, Garg K, Cheung CY, Larsen BB, Styrchak S, et al. HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science.* 2014;345(6196):570-3.
14. Simonetti FR, Sobolewski MD, Fyne E, Shao W, Spindler J, Hattori J, et al. Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo. *Proc Natl Acad Sci U S A.* 2016;113(7):1883-8.
15. von Andrian UH, and Mempel TR. Homing and cellular traffic in lymph nodes. *Nat Rev Immunol.* 2003;3(11):867-78.
16. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim EY, Archer J, Pond SLK, et al. Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature.* 2016;530(7588):51-6.
17. Schnell G, Joseph S, Spudich S, Price RW, and Swanstrom R. HIV-1 replication in the central nervous system occurs in two distinct cell types. *PLoS Pathog.* 2011;7(10):e1002286.
18. van Marle G, Gill MJ, Kolodka D, McManus L, Grant T, and Church DL. Compartmentalization of the gut viral reservoir in HIV-1 infected patients. *Retrovirology.* 2007;4:87.

19. Fletcher CV, Staskus K, Wietgrefe SW, Rothenberger M, Reilly C, Chipman JG, et al. Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues. *Proc Natl Acad Sci U S A.* 2014;111(6):2307-12.
20. Estes JD, Kityo C, Ssali F, Swainson L, Makamdop KN, Del Prete GQ, et al. Defining total-body AIDS-virus burden with implications for curative strategies. *Nature medicine.* 2017;23(11):1271.
21. Cory TJ, Schacker TW, Stevenson M, and Fletcher CV. Overcoming pharmacologic sanctuaries. *Current opinion in HIV and AIDS.* 2013;8(3):190-5.
22. Solas C, Lafeuillade A, Halfon P, Chadapaud S, Hittinger G, and Lacarelle B. Discrepancies between protease inhibitor concentrations and viral load in reservoirs and sanctuary sites in human immunodeficiency virus-infected patients. *Antimicrob Agents Chemother.* 2003;47(1):238-43.
23. Deeks SG, Kitchen CM, Liu L, Guo H, Gascon R, Narvaez AB, et al. Immune activation set point during early HIV infection predicts subsequent CD4+ T-cell changes independent of viral load. *Blood.* 2004;104(4):942-7.
24. Achaz G, Palmer S, Kearney M, Maldarelli F, Mellors JW, Coffin JM, et al. A robust measure of HIV-1 population turnover within chronically infected individuals. *Mol Biol Evol.* 2004;21(10):1902-12.
25. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, et al. Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat Med.* 2016;22(9):1043-9.
26. Coffin JM, Wells DW, Zerbato JM, Kuruc JD, Guo S, Luke BT, et al. Clones of infected cells arise early in HIV-infected individuals. *JCI Insight.* 2019;4(12).
27. Wiegand A, Spindler J, Hong FF, Shao W, Cyktor JC, Cillo AR, 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-e68.
28. Baxter AE, Niessl J, Fromentin R, Richard J, Porichis F, Charlebois R, et al. Single-Cell Characterization of Viral Translation-Competent Reservoirs in HIV-Infected Individuals. *Cell Host Microbe.* 2016;20(3):368-80.
29. Deleage C, Wietgrefe SW, Del Prete G, Morcock DR, Hao XP, Piatak M, Jr., et al. Defining HIV and SIV Reservoirs in Lymphoid Tissues. *Pathogens & immunity.* 2016;1(1):68-106.
30. Laird GM, Rosenbloom DI, Lai J, Siliciano RF, and Siliciano JD. Measuring the Frequency of Latent HIV-1 in Resting CD4(+) T Cells Using a Limiting Dilution Coculture Assay. *Methods in molecular biology (Clifton, NJ).* 2016;1354:239-53.
31. Hong F, Aga E, Cillo A, Yates AL, Besson G, Fyne E, et al. Novel assays to measure total cell-associated HIV-1 DNA and RNA. *J Clin Microbiol.* 2016.
32. Kearney M, Maldarelli F, Shao W, Margolick JB, Daar ES, Mellors JW, et al. Human immunodeficiency virus type 1 population genetics and adaptation in newly infected individuals. *J Virol.* 2009;83(6):2715-27.
33. Boritz EA, and Douek DC. Perspectives on Human Immunodeficiency Virus (HIV) Cure: HIV Persistence in Tissue. *J Infect Dis.* 2017;215(suppl_3):S128-S33.
34. Rainwater-Lovett K, Ziemniak C, Watson D, Luzuriaga K, Siberry G, Petru A, et al. Paucity of Intact Non-Induced Provirus with Early, Long-Term Antiretroviral Therapy of Perinatal HIV Infection. *PLoS One.* 2017;12(2):e0170548.
35. Ananworanich J, Chomont N, Fletcher JL, Pinyakorn S, Schuetz A, Sereti I, et al. Markers of HIV reservoir size and immune activation after treatment in acute HIV infection with and without raltegravir and maraviroc intensification. *J Virus Erad.* 2015;1(2):116-22.
36. Ananworanich J, Puthanakit T, Suntarattiwong P, Chokephaibulkit K, Kerr SJ, Fromentin R, et al. Reduced markers of HIV persistence and restricted HIV-specific immune responses after early antiretroviral therapy in children. *AIDS.* 2014;28(7):1015-20.

37. Crowell TA, Fletcher JL, Sereti I, Pinyakorn S, Dewar R, Krebs SJ, et al. Initiation of antiretroviral therapy before detection of colonic infiltration by HIV reduces viral reservoirs, inflammation and immune activation. *J Int AIDS Soc.* 2016;19(1):21163.
38. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell.* 2013;155(3):540-51.
39. Hosmane NN, Kwon KJ, Bruner KM, Capoferri AA, Beg S, Rosenbloom DI, 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):959-72.
40. Bui JK, Sobolewski MD, Keele BF, Spindler J, Musick A, Wiegand A, et al. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS pathogens.* 2017;13(3):e1006283.
41. Lorenzi JC, Cohen YZ, Cohn LB, Kreider EF, Barton JP, Learn GH, 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-E16.
42. Anderson EM, and Maldarelli F. Quantification of HIV DNA Using Droplet Digital PCR Techniques. *Curr Protoc Microbiol.* 2018;51(1):e62.
43. Palmer S, Wiegand AP, Maldarelli F, Bazmi H, Mican JM, Polis M, 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):4531-6.
44. Mansky LM, and Temin HM. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *Journal of virology.* 1995;69(8):5087-94.
45. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America.* 2008;105(21):7552-7.
46. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, 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):406-13.
47. Eypasch E, Lefering R, Kum CK, and Troidl H. Probability of adverse events that have not yet occurred: a statistical reminder. *BMJ.* 1995;311(7005):619-20.
48. Hudson RR, Boos DD, and Kaplan NL. A statistical test for detecting geographic subdivision. *Molecular biology and evolution.* 1992;9(1):138-51.
49. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, and Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology.* 2010;59(3):307-21.
50. Lefort V, Longueville JE, and Gascuel O. SMS: Smart Model Selection in PhyML. *Mol Biol Evol.* 2017;34(9):2422-4.
51. Critchlow DE, Li S, Nourijelyani K, and Pearl DK. Some statistical methods for phylogenetic trees with application to HIV disease. *Math Comput Model.* 2000;32(1-2):69-81.

Table 1: Infected Cell Clones in PBMC and LNMC

Patient Identifier (PID)	# of integration sites obtained[a]		Expanded clones detected[b]							
	PBMC	LNMC	Total # of clones detected	Clones detected in both PBMC and LNMC	Gene or nearest gene	Clones detected in PBMC only	Gene or nearest gene	Clones detected in LNMC only	Gene or nearest gene	p-value[c]
1683	143	72	9	5	*WIPF1, RAP1A, AATF, TFCP2, MAPK14*	3	*SIFLEC15, AGAP1, BRE*	1	*PDPK1*	
1079	174	63	7	3	*ACO1, TAF3, FBLN7*	4	*RNF213, JOSD1, COG4, ZNF490*	0	--	
2669	247	579	47	12	*NFATC3, CNN2, ZNF207, TPX2, FOPNL, XPO6, CUL3, CTSD, PPP4R1, SENP7, INPP4B, LOC374443*	10	*LOC729609, TFCP2, TMCO5A, SCAMP2, MARS, ZNF207, MARCH14, MKL1, USP4, PBRM1*	25	*TMX3, EPB41, PIBF1, FLOT1 DNAJC16, ARHGAP25, MIDN, LOC100505746, ASH1L, KDM2A, NUMA1, STAT5B, ARHGAP33, IFNAR2, FBXL17, TTF1, CCP110, XPO6, TRPV3, MARCH14, AP3D1, ACAP2, ABHD16A, RPS10- NUDT3, RAB39B*	0.1

[a]Integration sites assay (11)
[b]Clonal integration sites were identified ≥2 times in a single location or at least once in both locations
[c]Wilcoxon Signed Rank test to determine if differences in clonal detection is within sampling error

Table 2: Number of rakes of identical DNA sequences detected by SGS vs. number of clones of infected cells detected by ISA

Patient Identifier (PID)	Average P6-PR-RT, *env* pairwise distance (%)	# of rakes of identical P6-PR-RT sequences detected by SGS - PBMC and LNMC combined	# of rakes of identical *env* sequences detected by SGS - PBMC and LNMC combined	# of expanded clones[a] detected by integration sites assay – PBMC and LNMC combined
1683	0.5, 0.7	2	5	9
1079	2.4, 2.1	3	5	7
2669	2.1, 2.8	9	4	47

[a]Integration site detected more than once

Table 3: HIV-1 RNA levels in PBMC and LNMC

Treatment status	PID	Number of infected cells/million MC[a] (# proviruses with P6-PR-RT assayed)		Number infected cells/million CD4+ T cells[a]		Percent of infected cells with unspliced HIV RNA[a] (# cells with ca-HIV RNA assayed)		p-value	Mean number HIV RNA copies/infected cells (range)		p-value
		PBMC	LNMC	PBMC	LNMC	PBMC	LNMC		PBMC	LNMC	
Pre-ART	3720	233 (40)	994 (66)	3874	1614	10 (27)	100 (247)	-	4.6 (1-53)	2.2 (1-71)	-
	1683	837 (21)	ND[c]	2886	ND	13 (16)	ND		5.3 (1-57)	ND	
	1079	2677 (38)		14,089		8 (20)			2.1 (1-13)		
On ART	3720	6 (55)	ND (18)	34	ND	31 (18)	ND (3)	0.1[b]	2.6 (1-18)	1.3 (1-2)	>0.9[b]
	2661	6 (19)	<4 (25)	23	<8	9 (12)	<13 (12)		1.6 (1-4)	1.6 (1-3)	
	1683	192 (37)	154 (27)	708	423	5 (24)	7 (36)		1.6 (1-6)	1.3 (1-7)	
	1079	137 (32)	194 (40)	521	427	8 (103)	20 (138)		1.3 (1-13)	2.0 (1-16)	
	2669	429 (50)	1673 (97)	3226	3637	6 (262)	13 (108)		1.7 (1-12)	2.8 (1-30)	

[a]Integrase cell-associated DNA (iCAD) protocol (31)
[b]Wilcoxon signed-rank test
[c]Sample not available

Figure 1

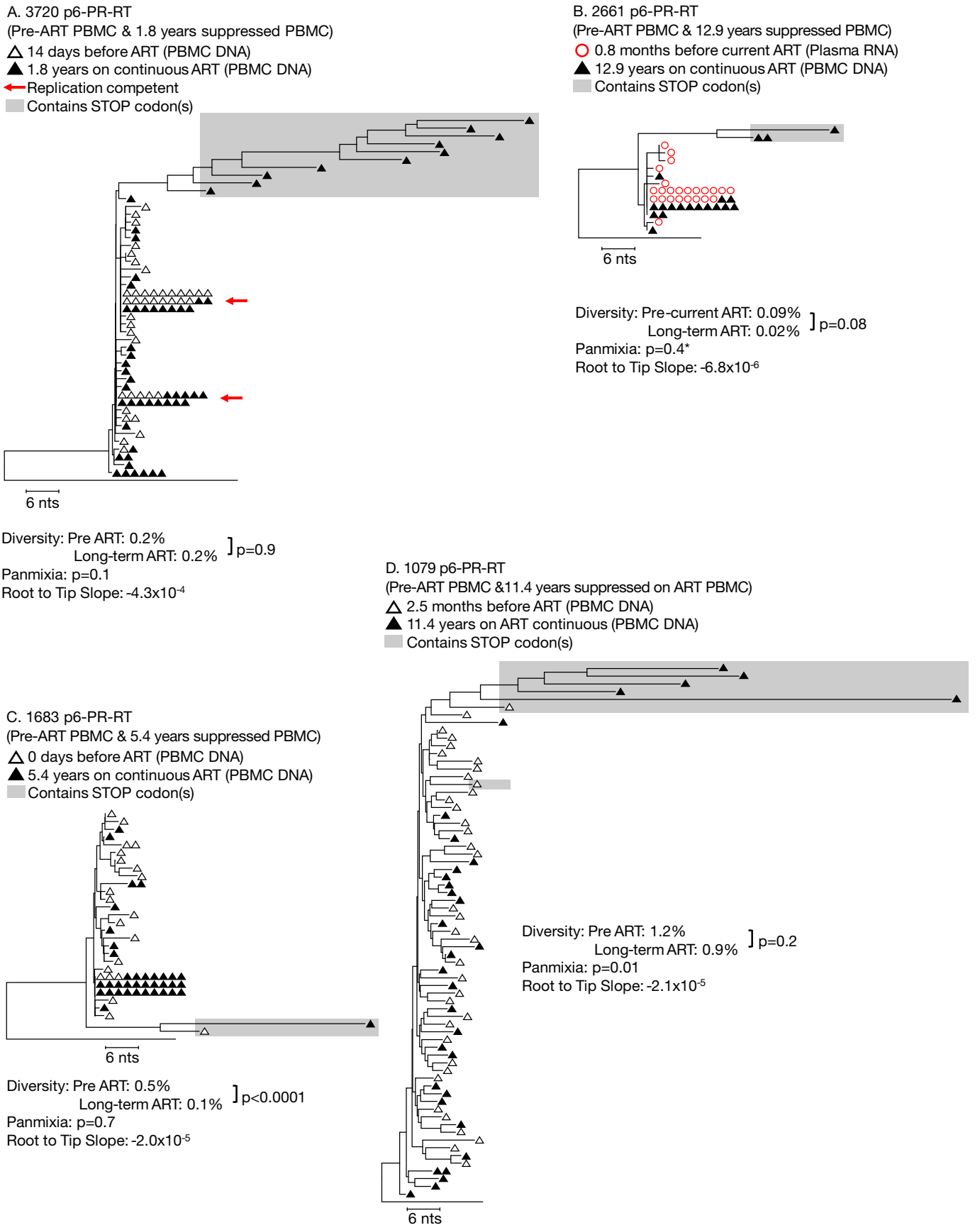
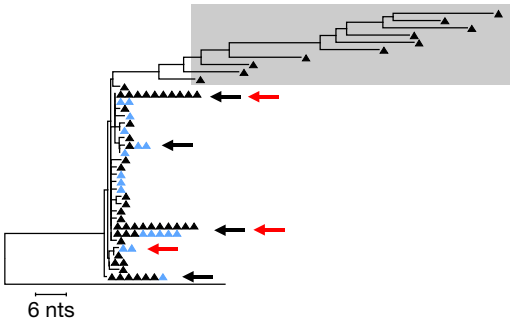

A. 3720 p6-PR-RT
(Pre-ART PBMC & 1.8 years suppressed PBMC)
14 days before ART (PBMC DNA)
1.8 years on continuous ART (PBMC DNA)
Replication competent
Contains STOP codon(s)
6 nts
Diversity: Pre ART: 0.2%
Long-term ART: 0.2%
p=0.9
Panmixia: p=0.1
Root to Tip Slope: -4.3x10-4
B. 2661 p6-PR-RT
(Pre-ART PBMC & 12.9 years suppressed PBMC)
0.8 months before current ART (Plasma RNA)
12.9 years on continuous ART (PBMC DNA)
Contains STOP codon(s)
6 nts
Diversity: Pre-current ART: 0.09%
Long-term ART: 0.02%
p=0.08
Panmixia: p=0.4*
Root to Tip Slope: -6.8x10-6
C. 1683 p6-PR-RT
(Pre-ART PBMC & 5.4 years suppressed PBMC)
0 days before ART (PBMC DNA)
5.4 years on continuous ART (PBMC DNA)
Contains STOP codon(s)
6 nts
Diversity: Pre ART: 0.5%
Long-term ART: 0.1%
p<0.0001
Panmixia: p=0.7
Root to Tip Slope: -2.0x10-5
D. 1079 p6-PR-RT
(Pre-ART PBMC &11.4 years suppressed on ART PBMC)
2.5 months before ART (PBMC DNA)
11.4 years on ART continuous (PBMC DNA)
Contains STOP codon(s)
6 nts
Diversity: Pre ART: 1.2%
Long-term ART: 0.9%
p=0.2
Panmixia: p=0.01
Root to Tip Slope: -2.1x10-5

Figure 1. HIV-1 P6-PR-RT proviral DNA and plasma virus RNA sequences prior to and during long-term ART. Neighbor joining trees were constructed from single-genome P6-PR-RT proviral DNA sequences obtained from PBMC (hollow black triangles; **A**, **C**, **D**) or RNA sequences obtained from plasma virus (hollow red circles; **B**) prior to the most recent period of continuous viral suppression on ART and proviral DNA sequences obtained from PBMC samples taken after 1.8-12.9 years of viral suppression on ART (solid black triangles). Diversity was measured by average pairwise distance (APD), and measures were compared using an unpaired t-test. Divergence was measured by a test for panmixia (to correct for multiple comparisons, populations are considered to be statistically different when the probability of panmixia is less than 0.001). Red arrows indicate sequences matching virus obtained in a viral outgrowth assay (30). Root-to-tip distances were measured using maximum likelihood trees, and the slopes of the root-to-tip distances over time were calculated by linear regression (units for slopes are substitutions/year). In cases where the slope was positive, an F-test was used to determine if the root-to-tip slopes were significantly different from zero. **A** is rooted on the HIV-1 subtype C consensus sequence, and **B**-**D** are rooted on the subtype B consensus sequence. Sequences containing G to A hypermutations and/or stop codons in open reading frames (indicated by shaded boxes) were excluded from all analyses. Except where unique variants were too few to test statistically (indicated by *), rakes of identical sequences were collapsed to a single variant for the test for panmixia and branch length analysis. Results from a total of 8 samples – two samples each from 4 patients – are represented in this figure.

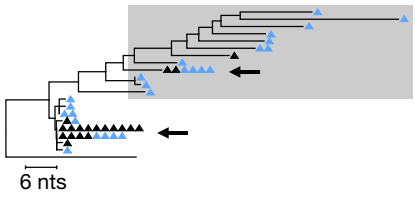
Figure 2

A. 3720 p6-PR-RT
(1.8 year suppressed PB vs. LN)



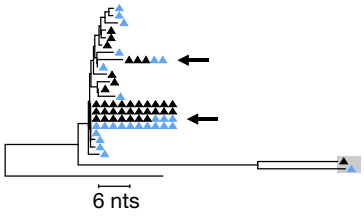
Diversity: PBMC: 0.2%
LNMC: 0.2%
p=0.7
Panmixia: p=0.9
Branch Length Correlation Coef.: 6.2×10^{-2}, p=0.2

B. 2661 p6-PR-RT
(12.9 years suppressed PB vs. LN)



Diversity: PBMC: 0.02%
LNMC: 0.1%
p=0.06
Panmixia: p=0.003*
Branch Length Correlation
Coef.: -6.5×10^{-2}, p=0.6*

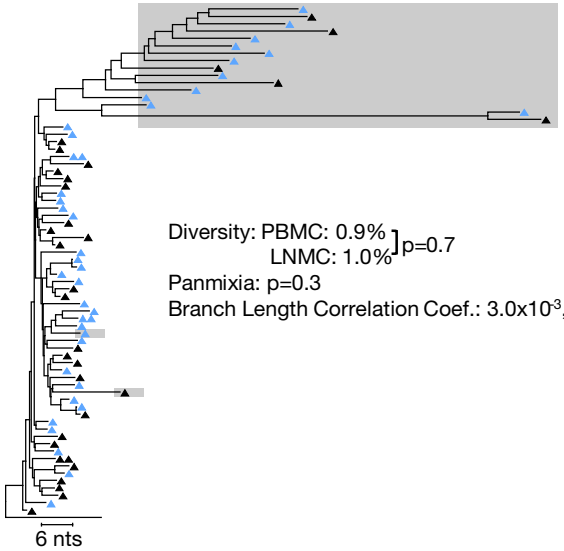
C. 1683 p6-PR-RT
(5.4 years suppressed PB vs. LN)



Diversity: PBMC: 0.1%
LNMC: 0.3%
p=0.04
Panmixia: p=0.8
Branch Length Correlation
Coef.: 4.1×10^{-2}, p=0.08

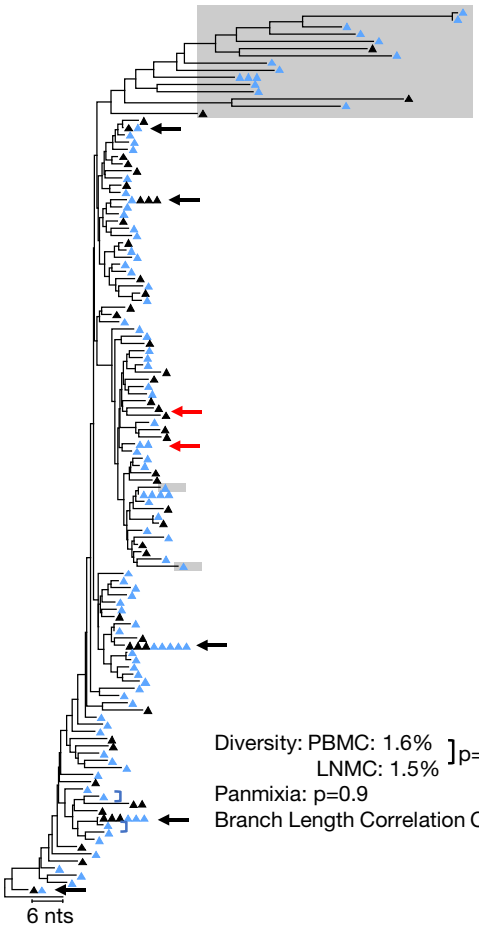
- PBMC HIV Provirus
- LNMC HIV Provirus
- Contains STOP codon(s)
- Clonal sequences
- Replication competent
- Contains NRTI and/or NNRTI resistance mutations

D. 1079 p6-PR-RT
(11.4 years suppressed PB vs. LN)



Diversity: PBMC: 0.9%
LNMC: 1.0%
p=0.7
Panmixia: p=0.3
Branch Length Correlation Coef.: 3.0×10^{-3}, p=0.3

E. 2669 p6-PR-RT
(4.3 years suppressed PB vs. LN)



Diversity: PBMC: 1.6%
LNMC: 1.5%
p=0.8
Panmixia: p=0.9
Branch Length Correlation Coef.: 2.7×10^{-3}, p=0.5

Figure 2. HIV-1 P6-PR-RT proviral DNA populations in lymph nodes and peripheral blood during ART. Neighbor joining trees were constructed from single-genome P6-PR-RT proviral sequences obtained from PBMC (black triangles) and LNMC (blue triangles) after 1.8-12.9 years of continuous viral suppression on ART. Diversity, divergence, and root-to-tip distances were measured as described in the legend to Figure 1. Branch length correlation coefficients were calculated from maximum likelihood trees, and significance was assessed by a permutation test as described in the methods. Black arrows indicate identical sequences that were found in both locations. The presence of identical viral sequences in both locations is likely due to clonal expansion of HIV-1 infected cells. Red arrows indicate sequences that matched virus recovered in the viral outgrowth assay. **A** is rooted on the HIV-1 subtype C consensus sequence; **B**-**E** are rooted on the HIV-1 subtype B consensus sequence. Sequences containing G to A hypermutation and/or stop codons in open reading frames (indicated by shaded boxes) were excluded from all analyses. Except where indicated (by *), rakes of identical sequences were collapsed to a single variant for the test for panmixia and branch length analysis. Results from a total of 10 samples – two samples each from 5 patients – are represented in this figure.

Figure 3

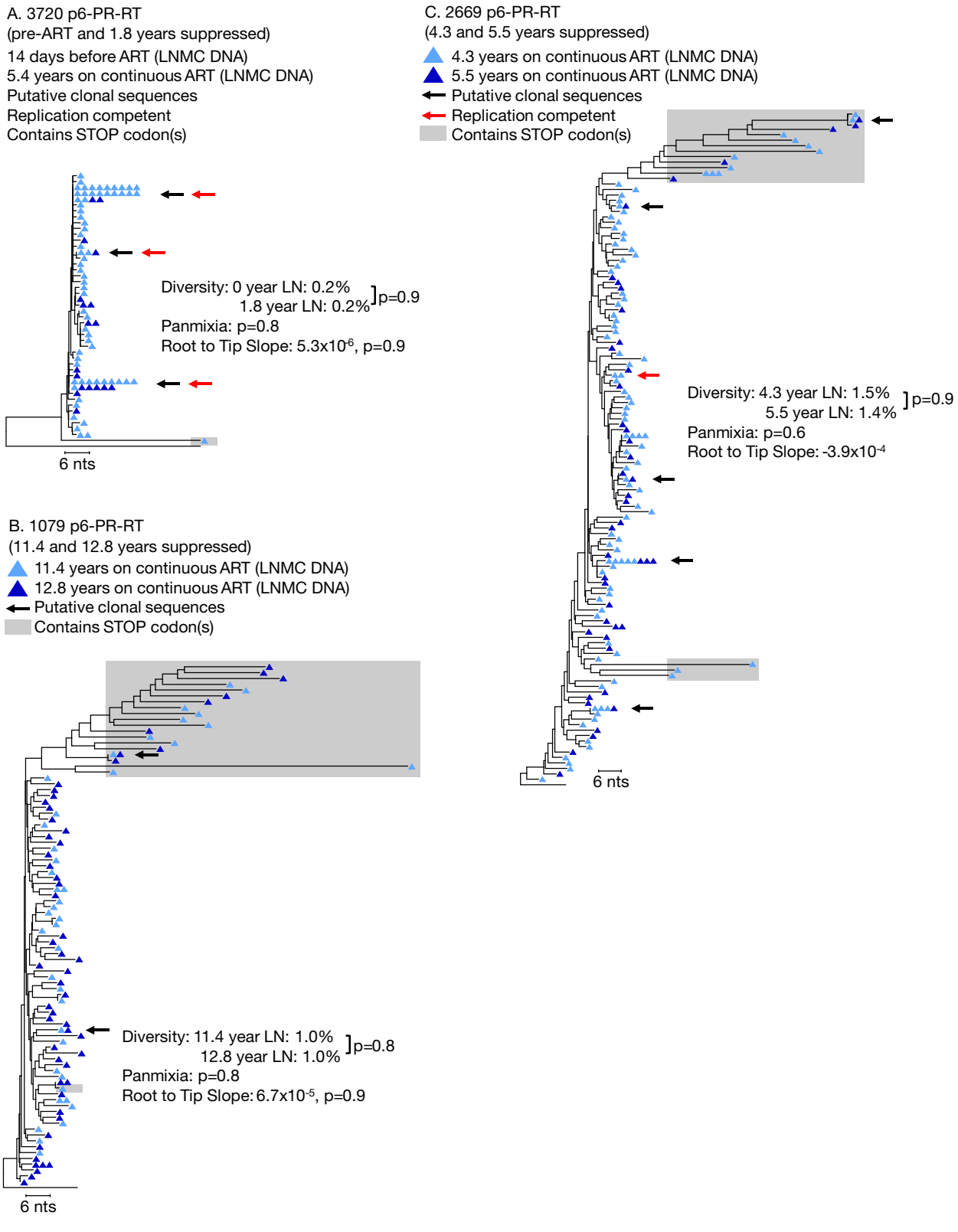

A. 3720 p6-PR-RT
(pre-ART and 1.8 years suppressed)
14 days before ART (LNMC DNA)
5.4 years on continuous ART (LNMC DNA)
Putative clonal sequences
Replication competent
Contains STOP codon(s)
Diversity: 0 year LN: 0.2%
1.8 year LN: 0.2%
p=0.9
Panmixia: p=0.8
Root to Tip Slope: 5.3x10-6, p=0.9
6 nts
B. 1079 p6-PR-RT
(11.4 and 12.8 years suppressed)
11.4 years on continuous ART (LNMC DNA)
12.8 years on continuous ART (LNMC DNA)
Putative clonal sequences
Contains STOP codon(s)
Diversity: 11.4 year LN: 1.0%
12.8 year LN: 1.0%
p=0.8
Panmixia: p=0.8
Root to Tip Slope: 6.7x10-5, p=0.9
6 nts
C. 2669 p6-PR-RT
(4.3 and 5.5 years suppressed)
4.3 years on continuous ART (LNMC DNA)
5.5 years on continuous ART (LNMC DNA)
Putative clonal sequences
Replication competent
Contains STOP codon(s)
Diversity: 4.3 year LN: 1.5%
5.5 year LN: 1.4%
p=0.9
Panmixia: p=0.6
Root to Tip Slope: -3.9x10-4
6 nts

Figure 3. HIV-1 P6-PR-RT proviral DNA sequences from longitudinal lymph node samples. Neighbor joining trees were constructed from single-genome P6-PR-RT proviral sequences obtained from LNMC at two time points: the first prior to ART initiation or after 11.4 or 4.3 years on ART (dark blue triangles) and the second approximately one year later (light blue triangles). Diversity, divergence, and root-to-tip distances were measured as described in the legend to Figure 1. Black arrows indicate clonal sequences present at both time points, and red arrows indicate sequences matching virus that grew in the viral outgrowth assay. **A** is rooted on the HIV-1 subtype C consensus sequence, and **B** and **C** are rooted on the subtype B consensus sequence. Sequences containing G to A hypermutation and/or stop codons in open reading frames (indicated by shaded boxes) were excluded from all analyses. Results from a total of 6 samples – two samples each from 3 patients – are represented in this figure.

Figure 4

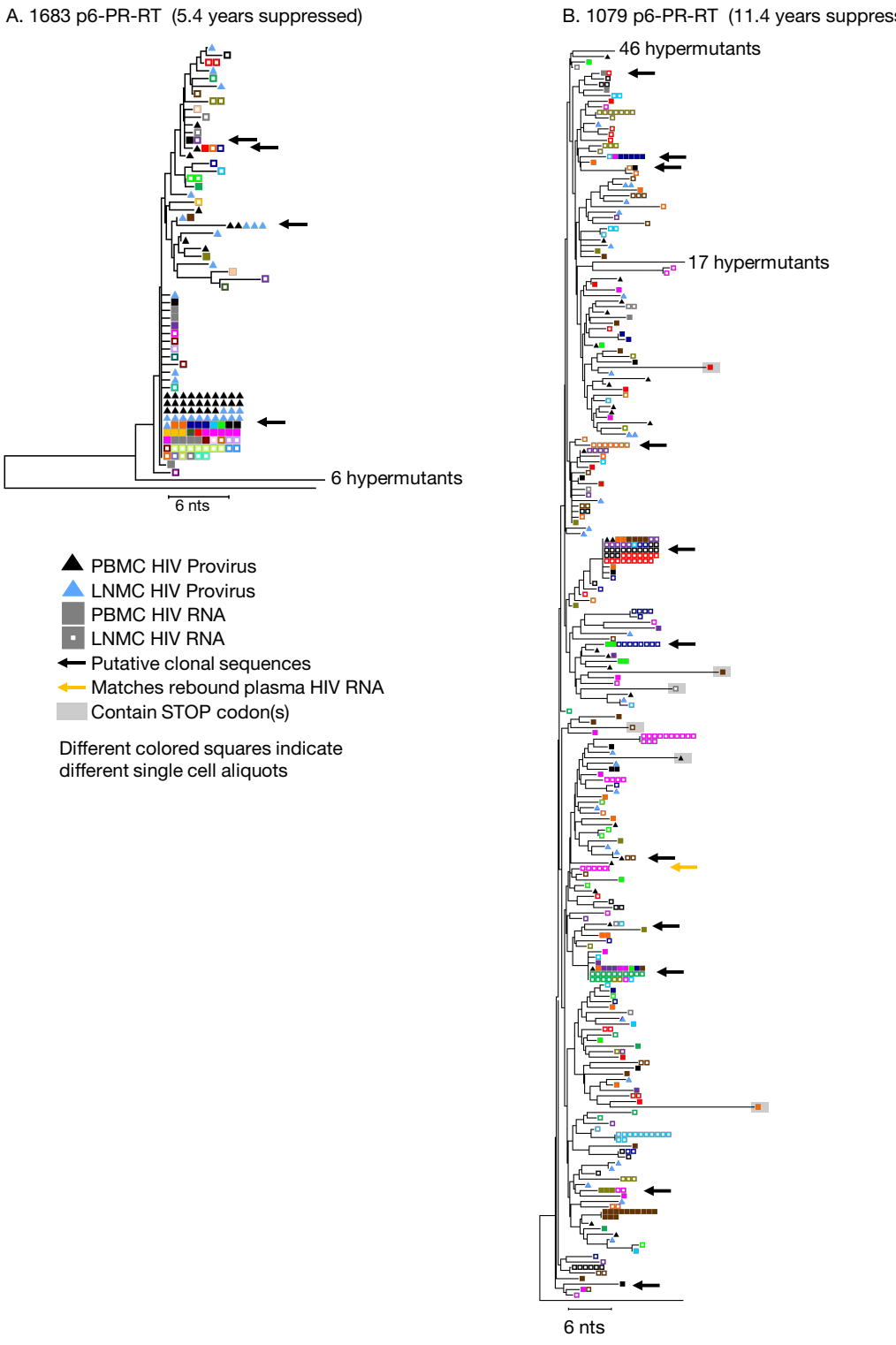

A. 1683 p6-PR-RT (5.4 years suppressed)
6 hypermutants
6 nts
PBMC HIV Provirus
LNMC HIV Provirus
PBMC HIV RNA
LNMC HIV RNA
Putative clonal sequences
Matches rebound plasma HIV RNA
Contain STOP codon(s)
Different colored squares indicate
different single cell aliquots
B. 1079 p6-PR-RT (11.4 years suppressed)
46 hypermutants
17 hypermutants
6 nts

Figure 4. HIV-1 cell-associated RNA sequences obtained from single cells in peripheral blood and in lymph nodes during ART. Neighbor joining trees were constructed from single-genome P6-PR-RT proviral DNA and cell-associated HIV-1 RNA sequences obtained from paired PBMC and LNMC sampled during continuous viral suppression on ART. Black and blue triangles indicate HIV-1 DNA sequences from PBMC and LNMC, respectively; solid squares and hollow squares indicate HIV-1 RNA sequences from PBMC and LNMC, respectively. Squares with no genetic difference of the same color indicate sequences that are assumed to be from the expression of single infected cells. Black arrows indicate sequences detected in probable clones in both the samples. The yellow arrow indicates a sequence detected in rebound plasma viremia. Trees are rooted on the subtype B consensus sequence. Results from a total of 8 samples – four samples each from 2 patients – are represented in this figure.

Figure 5

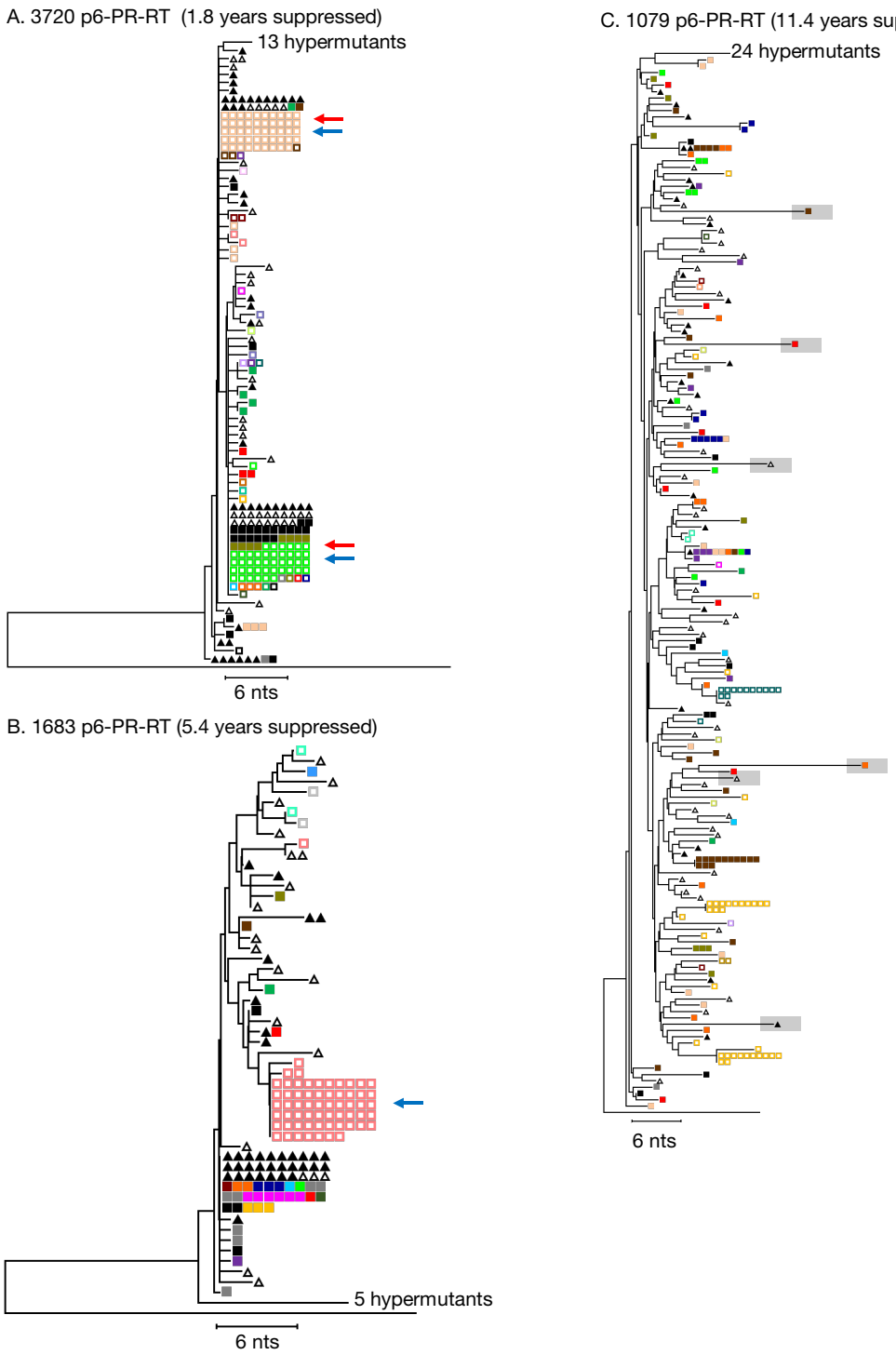

A. 3720 p6-PR-RT (1.8 years suppressed)
13 hypermutants
6 nts
B. 1683 p6-PR-RT (5.4 years suppressed)
5 hypermutants
6 nts
C. 1079 p6-PR-RT (11.4 years suppressed)
24 hypermutants
6 nts
Pre-ART PBMC HIV provirus
On-ART PBMC HIV provirus
On-ART PBMC HIV RNA
Pre-ART PBMC HIV RNA
High expressing cells
Replication competent
Contains STOP codon(s)
Different colored squares indicate different aliquots with few cells with HIV RNA

Figure 5. HIV-1 cell-associated RNA from single cells in peripheral blood prior to and during long-term ART. Neighbor joining trees were constructed from single-genome P6-PR-RT proviral DNA and cell-associated HIV-1 RNA sequences obtained from PBMC sampled prior to ART initiation (hollow triangles indicate proviral sequences, hollow squares indicate RNA sequences) and after 1.8-11.4 years of viral suppression on ART (solid triangles indicate proviral sequences, solid squares indicate RNA sequences). Squares of the same color indicate sequences that are assumed to be from the expression of single infected cells. The blue arrows indicate high expressing cells (>20 HIV-1 RNA copies), and red arrows indicate sequences matching virus that grew in the viral outgrowth assay. **A** is rooted on the HIV-1 subtype C consensus sequence, and **B** and **C** are rooted on the subtype B consensus sequence. Results from a total of 12 samples – four samples each from 3 patients – are represented in this figure.

Figure 6

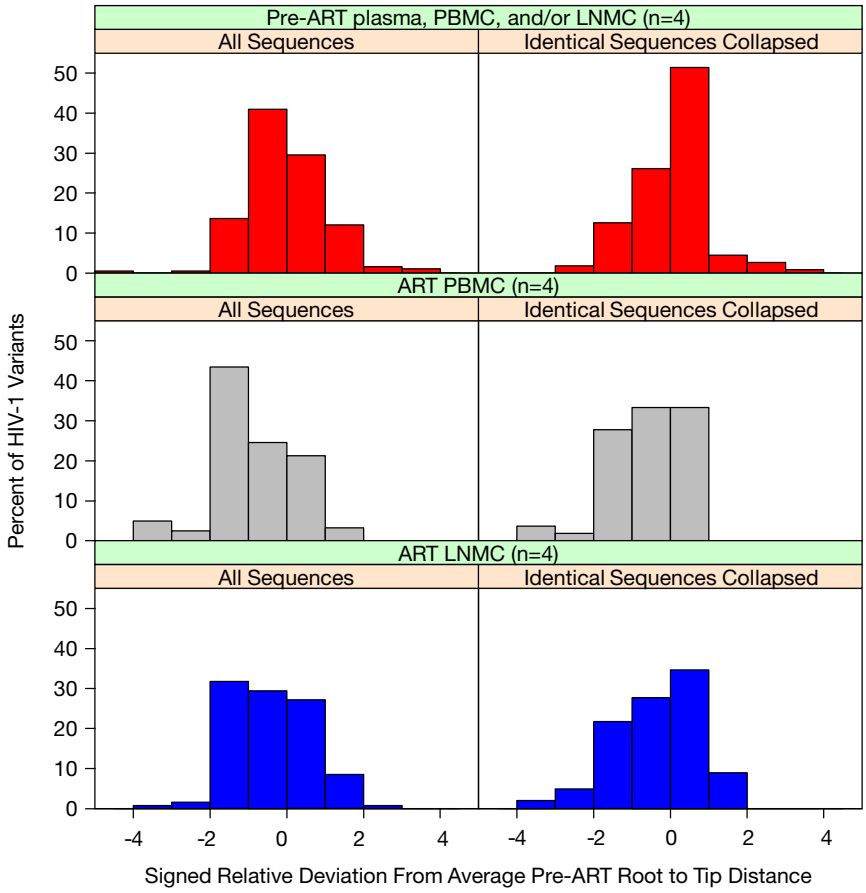

Pre-ART plasma, PBMC, and/or LNMC (n=4)
All Sequences
Identical Sequences Collapsed
ART PBMC (n=4)
All Sequences
Identical Sequences Collapsed
ART LNMC (n=4)
All Sequences
Identical Sequences Collapsed
Percent of HIV-1 Variants
0
10
20
30
40
50
-4
-2
0
2
4
Signed Relative Deviation From Average Pre-ART Root to Tip Distance

Figure 6. HIV-1 root-to-tip distances normalized to pre-ART average. Signed relative standard deviations were calculated for each sequence (including either all sequences or only each distinct variant) from each patient, as described in the methods. These data were separated into categories of pre-ART, on-ART PBMC, and on-ART LNMC and the values from all patients were aggregated and plotted separately for each group in red, grey, and blue for pre-ART, on-ART PBMC, and on-ART LNMC, respectively, in bin sizes of 1 standard deviation. A negative relative standard deviation indicates a regression from the mean and a positive value indicates an increase from the mean. Patient 2669 was excluded from this analysis due to lack of pre-ART sampling. Sequences from patient 2661 included pre-ART plasma sequence data, and sequences from patient 3720 included both pre-ART PBMC and LNMC data. All sequences containing G to A hypermutation and/or stop codons in open reading frames were excluded from the analysis. The number of samples for each panel was 4.