Longitudinal clonal dynamics of HIV-1 latent reservoirs measured by combination quadruplex polymerase chain reaction and sequencing

Alice Cho[a], Christian Gaebler[a], Thiago Olveira[a], Victor Ramos[a], Marwa Saad[a], Julio C. C. Lorenzi[a], Anna Gazumyan[a], Susan Moir[b], Marina Caskey[a], Tae-Wook Chun[b], and Michel C. Nussenzweig[a,c,1]

[a]Laboratory of Molecular Immunology, The Rockefeller University, New York, NY 10065; [b]Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892; and [c]HHMI, The Rockefeller University, New York, NY 10065

Contributed by Michel C. Nussenzweig; received September 27, 2021; accepted December 14, 2021; reviewed by Andres Finzi and Una O'Doherty

HIV-1 infection produces a long-lived reservoir of latently infected CD4$^+$ T cells that represents the major barrier to HIV-1 cure. The reservoir contains both intact and defective proviruses, but only the proviruses that are intact can reinitiate infection upon cessation of antiretroviral therapy (ART). Here we combine four-color quantitative PCR and next-generation sequencing (Q4PCR) to distinguish intact and defective proviruses and measure reservoir content longitudinally in 12 infected individuals. Q4PCR differs from other PCR-based methods in that the amplified proviruses are sequence verified as intact or defective. Samples were collected systematically over the course of up to 10 y beginning shortly after the initiation of ART. The size of the defective reservoir was relatively stable with minimal decay during the 10-y observation period. In contrast, the intact proviral reservoir decayed with an estimated half-life of 4.9 y. Nevertheless, both intact and defective proviral reservoirs are dynamic. As a result, the fraction of intact proviruses found in expanded clones of CD4$^+$ T cells increases over time with a concomitant decrease in overall reservoir complexity. Thus, reservoir decay measurements by Q4PCR are quantitatively similar to viral outgrowth assay (VOA) and intact proviral DNA PCR assay (IPDA) with the addition of sequence information that distinguishes intact and defective proviruses and informs reservoir dynamics. The data are consistent with the notion that intact and defective proviruses are under distinct selective pressure, and that the intact proviral reservoir is progressively enriched in expanded clones of CD4$^+$ T cells resulting in diminishing complexity over time.

HIV | latency | reservoir | Q4PCR

HIV is a major global health problem, with an estimated 38 million people living with HIV, leading to over 30 million HIV-related deaths as of 2020 (1, 2). While treatment with antiretroviral therapy (ART) effectively suppresses HIV-1 replication and reduces viremia (3), it does not cure infection. Treatment interruption typically leads to viral rebound within 2 to 4 wk, emanating from a reservoir of intact transcriptionally silent proviruses that persist in the genome of CD4$^+$ T cells (4–6). Thus, the intact proviral reservoir is the major barrier to HIV-1 cure (7–12).

Several different methods have been used to measure the HIV-1 latent reservoir and estimate its half-life, many of which cannot distinguish between intact and defective proviruses (9, 13, 14). Interpreting the results of these measurements is complicated by the observation that the great majority of the proviruses integrated into the genome of CD4$^+$ T cells are defective and that the decay rates for the two types of proviruses are different (13–19). For example, PCR measurements that rely on individual oligonucleotide probes do not distinguish between intact and defective proviruses, and therefore any measure of HIV-1 decay by such methods is primarily a measure of the defective reservoir (20). Quantitative viral outgrowth assays (QVOAs) are limited to the intact proviral reservoir and underestimate the size of the reservoir because only a fraction of the latent cells in the cultures can be activated by stimulation in vitro (10, 21–23). Nevertheless, assuming that this fraction is stable, QVOA should be a relatively accurate measure of the rate of reservoir decay (21). The newly developed intact proviral DNA assay (IPDA) is a high-throughput assay that uses two sets of oligonucleotide probes to regions in the HIV genome that are conserved across most clade B viruses (24). However, IPDA does not verify whether provirus sequences that hybridize to the two probes are fully intact (15, 24). As a result, IPDA may overestimate the size of the reservoir, and the accuracy of intact reservoir decay measurements may be variably altered by contaminating defective proviruses, especially in individuals with small intact and large defective reservoirs (20). In addition, IPDA cannot be used to examine reservoir clonal dynamics. Q4PCR combines a four-probe quantitative PCR strategy with near full-length amplification and sequence verification of intact or defective genomes (25). However, Q4PCR is a low-throughput assay, and the requirement for full-length genome amplification renders it less efficient than IPDA (20). Moreover, Q4PCR captures only the subset of defective proviruses that are detected by at least two of the four oligonucleotide probe sets (25).

To determine whether Q4PCR can be used to examine changes in the reservoir over time and document longitudinal reservoir clonal dynamics, we sampled a cohort of 12 people living with

Significance

HIV-1 infection requires lifelong treatment with antiretroviral therapy (ART) due to viral rebound of a latent reservoir of intact, transcriptionally silent provirus found to persist in the genome of CD4$^+$ T cells. One major challenge to understanding the nature of the latent reservoir is accurately characterizing the measuring of the size of the reservoir. Herein, we use quadruplex polymerase chain reaction (Q4PCR) to assess the dynamics of the latent reservoir in HIV$^+$ individuals who have been on long-term ART for up to 10 y. Our results show that Q4PCR can be used to accurately measure the latent reservoir, while providing the added benefit of assessing the genetic diversity of the reservoir to better understand changes to clonal dynamics over time.

Author contributions: A.C., C.G., M.C., and M.C.N. designed research; A.C., M.S., J.C.C.L., S.M., and T.-W.C. performed research; A.C., C.G., T.O., V.R., M.C., and M.C.N. analyzed data; A.G. contributed new reagents/analytic tools; S.M. and T.-W.C. recruited patients and executed clinical protocols; and A.C. and M.C.N. wrote the paper.

Reviewers: A.F., Université de Montréal; and U.O., Perelman School of Medicine, University of Pennsylvania.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: nussen@rockefeller.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2117630119/-/DCSupplemental.

Published January 18, 2022.

HIV over a period of 10 y starting shortly after ART initiation. Reservoir half-lives measured by Q4PCR were consistent with those reported for QVOAs and IPDA during equivalent observation intervals (10, 11, 17, 19). As reported for IPDA (17), the decay rates of intact and defective proviral reservoirs differed significantly. Sequencing revealed that clonality increases and diversity decreases over time in both the intact and defective proviral reservoirs.

Results

Characteristics of Study Participants. To investigate changes in the latent reservoir over time in people living with HIV on long-term ART, we assayed serum and peripheral blood mononuclear cell (PBMC) samples collected from 12 individuals who were treated with ART (*SI Appendix*, Table S1).

PBMC samples were collected at three or four time points: 1 to 9 mo; 11 to 15 mo; 2 to 3 y; and 5 or 10 y after initiating ART (Fig. 1*A*). For three individuals, additional plasma samples collected 1 to 6 mo before starting ART were used to compare circulating viral genomes to proviruses in the latent reservoir (Fig. 1*A*). All individuals had viral loads below 50 copies/mL at the evaluated time points, except for participant P7, who showed a minor viral blip at 59 copies/mL at the 2- to 3-y time point before returning to below 50.

Decay Rate of Intact and Defective Proviral Reservoir. Sequence information was obtained for 323 intact and 4,915 defective proviruses, with an average of over 100 sequences per participant per time point assayed (range: 9 to 355 sequences). On average, 93% of the latent reservoir sampled by Q4PCR comprised defective proviruses at all time points assayed (Fig. 1*B*). For two participants, no intact sequences were detected at any time point, and an additional four participants had a single time point where no intact proviruses were detected (see *Materials and Methods* for additional information).

Individuals showed variable intact and defective reservoir dynamics (*SI Appendix*, Fig. S1). Nevertheless, when considered together there were significant differences between the two reservoir compartments. Whereas the intact proviral reservoir decayed with a half-life of 4.9 y (Fig. 1*C*, $R = -0.64$, $P < 0.0001$), the defective proviral reservoir was far more stable, with an estimated half-life of 50 y (Fig. 1*D*). As a result, there was a significant proportional increase in defective proviruses from 90% of the total HIV-1 reservoir at 1 to 9 mo to 98% of sequences at the 5- to 10-y time point ($P = 0.009$, Fig. 1*B*). We conclude that the half-life of the intact and defective proviral reservoirs in this cohort measured by Q4PCR differ, and that they are entirely consistent with the measurements reported by QVOA and IPDA in larger studies (11, 15, 17, 21).

Reservoir Evolution during Long-Term ART. To examine reservoir sequence evolution over time, we compared circulating viral envelope (*env*) sequences obtained from plasma of three participants before they started ART with their corresponding proviral reservoir sequences from time points after ART initiation (Fig. 1*A*). Although we did not find identical sequences between the two compartments, plasma and reservoir sequences clustered closely together (Fig. 2*A*).

To further study the phylogenetic relationship between viruses circulating before ART initiation and proviruses in the reservoir, we compared their average patristic distances. As reported by others (26), the average distance did not change over time in any of the participants whether we considered all *env* sequences or only intact proviruses (Fig. 2*B* and *SI Appendix*, Fig. S2). Similar results were obtained when comparing the patristic distance between latent proviruses and their closest pre-ART genetic

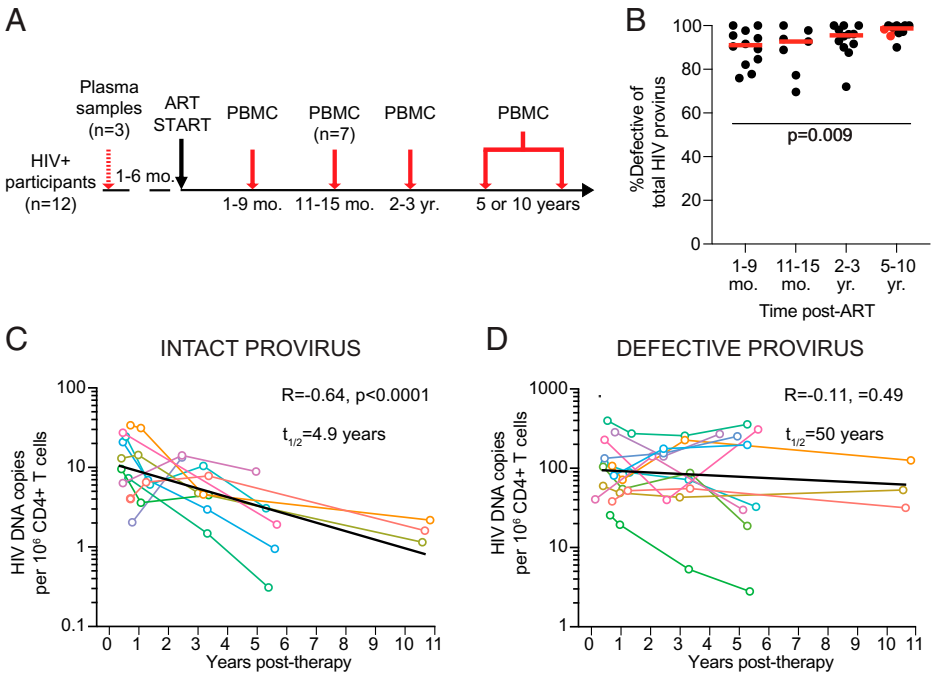

Fig. 1. Longitudinal measure of the HIV-1 provirus latent reservoir by Q4PCR. (*A*) Study design. PBMC samples were collected from all 12 participants at three to four time points after starting ART, with plasma samples acquired from a subset ($n = 3$) of participants prior to treatment. (*B*) Relative frequency of the defective HIV-1 proviral reservoir over time post-ART. Horizontal red line indicates median frequency. Participants sampled at 10 y post-ART are highlighted in red. (*C*) Number of intact HIV-1 proviruses per million $CD4^+$ T cells detected overt time. Half-life of intact provirus reservoir was calculated to be 4.9 y ($R = -0.64$, $P < 0.0001$). (*D*) Number of defective HIV-1 proviruses per million $CD4^+$ T cells detected over time. The half-life of the defective provirus reservoir was measured to be 50 y ($R = -0.11$, $P = 0.49$). Each colored line represents one individual participant. The exponential decay half-life was calculated by using log10 infectious units per million (IUPM) values and a linear regression model. Black line represents estimated decay of total dataset. A Kruskal–Wallis test with subsequent Dunn's multiple comparisons test was used to analyze data where appropriate.

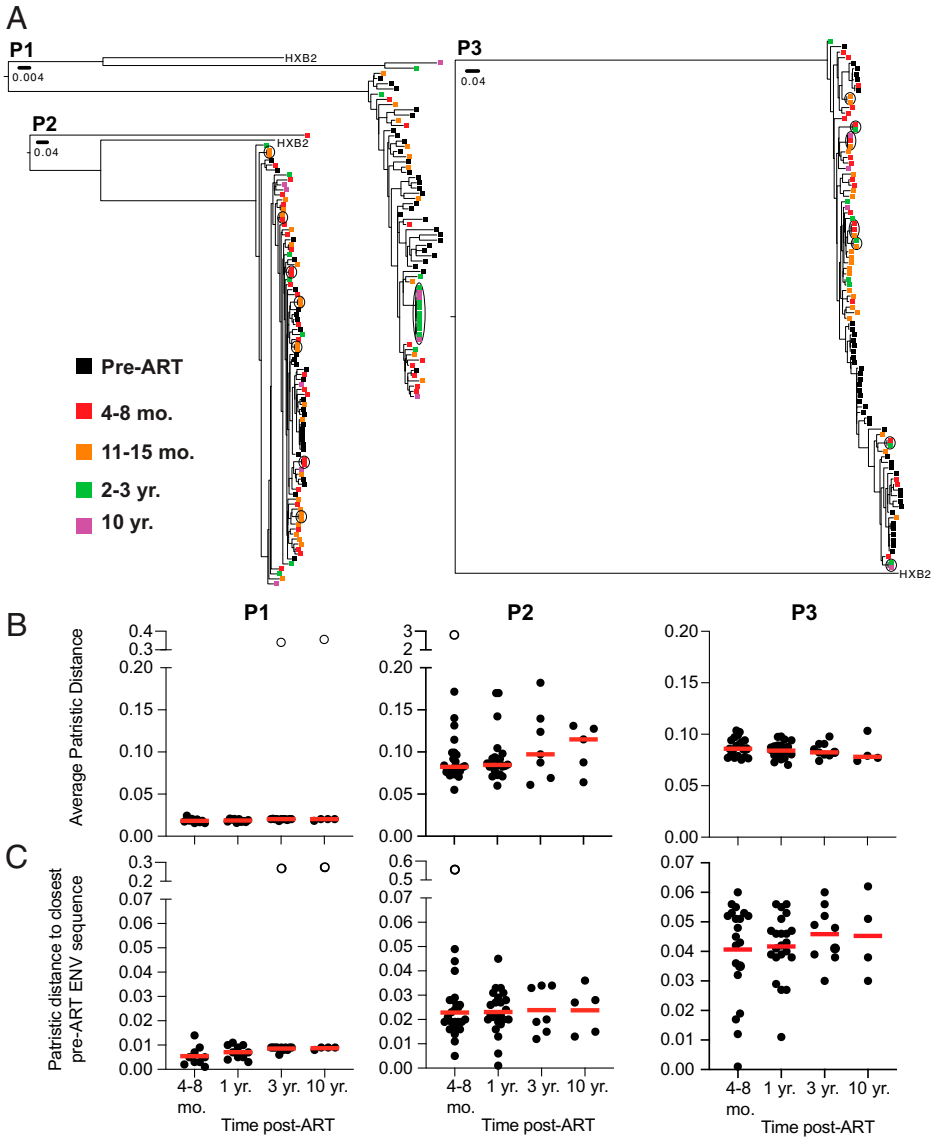

Fig. 2. Proviruses detected in circulation prior to ART are distinct from intact proviruses detected in latent reservoir found in PBMCs. (*A*) Phylogenetic tree of *env* sequences of intact proviruses in the latent reservoir and pre-ART proviruses found in circulation found in participants P1–P3. Branch lengths are proportional to a genetic distance (Scale bar indicated below sample ID). Black outlines indicate clones. Color of squares represent the origin of the sequence, as indicated by the key on the *Left*. (*B*) Graph showing average patristic distance between each intact provirus found in the reservoir compared to every pre-ART virus found in circulation, comparing different time points post-ART. Each dot represents one intact provirus found in the reservoir. (*C*) Graph showing distance of intact proviruses found in the latent reservoir to the closest pre-ART *env* sequence. Each dot represents one intact provirus found in the reservoir. Participant name is specified above graphs. Unfilled dots indicate outliers that were excluded from the analysis. Horizontal red line indicates the median patristic distance at each time point. Grubb's test was used to calculate outliers, and a Kruskal–Wallis test with subsequent Dunn's multiple comparisons was used to analyze data where appropriate.

relatives (Fig. 2*C*). However, some of the intact proviruses in the early reservoir that were most closely related to circulating viruses disappeared, confirming that many of the changes in the reservoir over time represent concomitant loss of variants and clonal proliferation, rather than diversification because of ongoing viral replication.

Clonal Dynamics in the Defective Proviral Reservoir. Neither QVOA nor IPDA examines the genetic diversity of the reservoir (11, 17, 21, 24). As expected from the disproportionally greater number of defective proviruses, we retrieved more defective than intact proviruses at all time points: 1,350 at 1 to 9 mo; 728 sequences at 11 to 15 mo; 1,214 sequences after 2 to 3 y; and 1,623 sequences after 5 to 10 y (Fig. 3*A*). The lower number of sequences obtained at the 11- to 15-mo time point reflects the smaller number of individuals sampled at that time point (Fig. 1*A*).

The overall representation of $CD4^+$ T cell clones that contain integrated defective proviruses differs between individuals but increases over time ($P = 0.0004$, Fig. 3*B*). When defective proviruses were considered in aggregate, only 16% of all sequences were found in clones at the early time point and this increases to 46% after 5 to 10 y ($P < 0.0001$, Fig. 3*C*). The defective reservoir was dynamic, with only 3% of the clonally expanded proviruses persisting throughout the observation period (Fig. 3*C*). Clones found uniquely at one time point constitute 5%, 3%, 7%, and 17% of the overall reservoir at each of the four time points, respectively (Fig. 3*C*). To determine how the increase in clonality may impact overall diversity, we used the Simpson index to determine the probability that two randomly chosen proviruses are

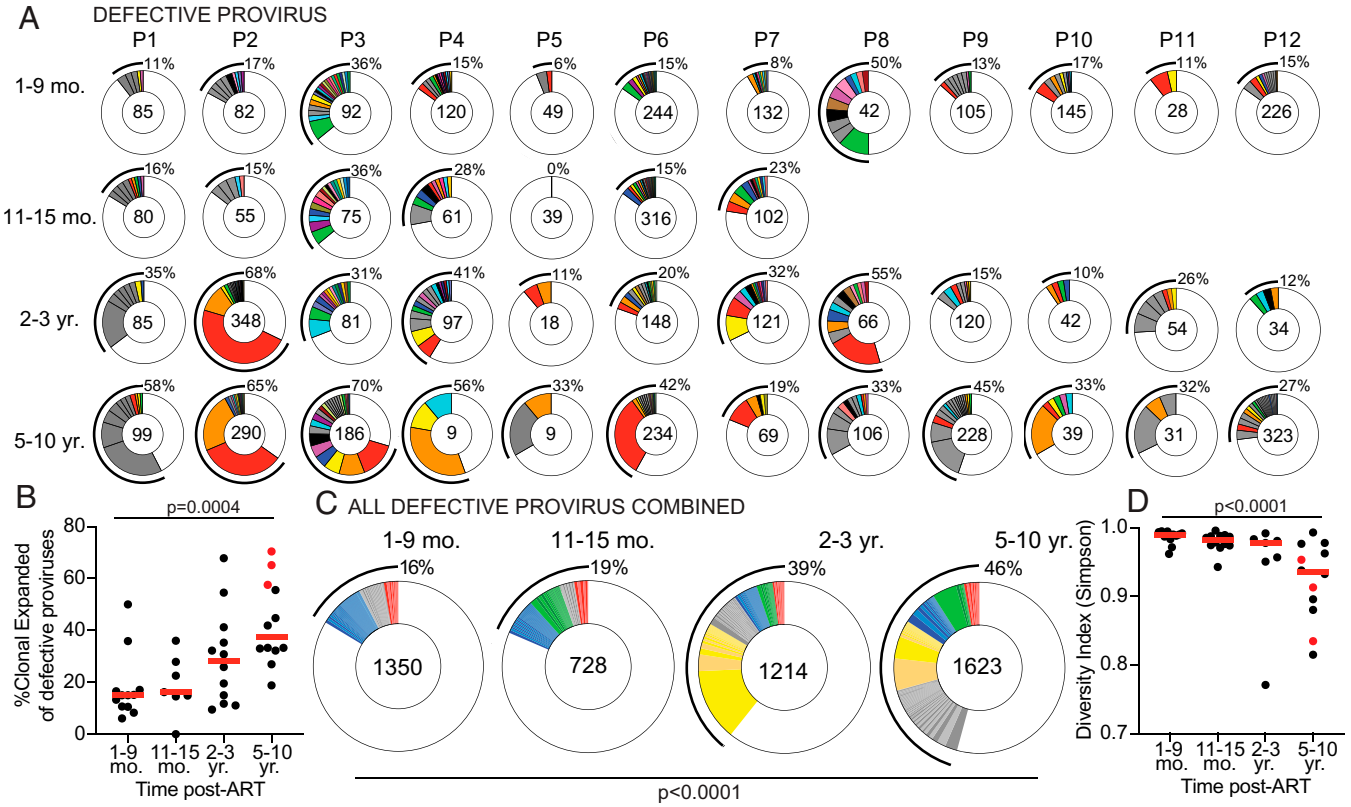

Fig. 3. Clonal dynamics of defective provirus reservoir over time. (*A*) Pie charts of the defective provirus reservoir detected in each individual, as denoted above the circles. Number inside each circle indicates the number of defective proviruses analyzed for each time point, specified to the *Left* of the circles. Color slices indicate persisting defective clones found at multiple time points, while gray slices are clonal expansions unique to the time point. Remaining white slices represent sequences isolated only once. Black outlines indicate the frequency of clonally expanded defective proviruses detected in each participant. (*B*) Graphs showing frequency of clonally expanded defective provirus detected in participants over time. Each dot represents one individual. Participants sampled at 10 y post-ART are highlighted in red. Horizontal red bars indicate the median frequency of clonality at each time point. (*C*) Pie charts showing persisting or expanded clones of all defective proviruses from all individuals combined, organized by when the clone was first detected. Time point assayed is denoted above the circles. Number inside each circle indicates the number of defective proviruses analyzed for each time point. Color slices indicate the following: gray, clonally expanded defective proviruses unique to the time point assays; blue, persisting clones of defective provirus, first detected at 1 to 9 mo post-ART; green, persisting clones first detected at 11 to 15 mo post-ART; yellow, persisting clones first detected 2 to 3 y post-ART; red, persisting clones detected at all assayed time points; and white, defective provirus sequences isolated only once. Black line indicates the frequency of clonally expanded defective proviruses detected at each time point. (*D*) Graph showing diversity (determined by Simpson's index) of all proviruses detected in participants over time. Each dot represents one individual. Participants sampled at 10 y post-ART are highlighted in red. Horizontal red bars indicate the median frequency of clonality at each time point. A Kruskal–Wallis test with subsequent Dunn's multiple comparisons was used to analyze data where appropriate. A Fisher's exact test with subsequent Bonferroni correction was used to compare clonality of defective proviruses between time points.

from the same clone (Fig. 3*D*). From this analysis, we see that the diversity significantly decreases over time ($P < 0.0001$, Fig. 3*D*). Thus, the increased clonal representation results in a decrease of the overall proviral complexity in the reservoir.

The defective reservoir is composed of proviral genomes that contain undamaged open reading frames (ORFs) with the potential to encode HIV-1 proteins and others that cannot (27–30). For example, whereas proviruses with smaller internal deletions may be able to splice together mRNAs that encode some HIV-1 proteins, those with larger deletions, hypermutations, or mutations in the major splice donor (MSD) are less likely to be translation competent and produce proteins (28, 31). It has been suggested that immune selective pressure, and therefore the kinetics of decay, differs between defective proviruses that contain ORFs and can produce proteins and those that cannot (28, 31).

To determine whether the decay kinetics of different types of defective proviruses are heterogeneous, we separated defective proviral sequences into five categories: 1) Missing internal genes but able to encode protein products; 2) noncoding that contain at least one early stop codon but possibly able to encode some functional proteins; 3) MSD mutation unable to produce spliced mRNA; 4) inversions or duplications; and 5) uncategorized defective proviruses (*SI Appendix*, Table S2). All the different categories of defective proviruses show increased frequencies of clonally expanded sequences over time (Fig. 4 *A*–*G*, and *SI Appendix*, Fig. S3). While overall clonality increased, there was variability in how specific clones wax and wane over time, with some defective proviruses showing large expansions at later time points (Fig. 4 *A*–*E*). One possibility is that these differences reflect ongoing responses to viral or other infections (32, 33). Consistent with the idea that defective proviruses that contain undamaged ORFs are under different selective pressure than those that are unlikely to be translation competent, proviruses missing internal genes decayed faster than noncoding proviruses (Fig. 4 *H* and *I* and *SI Appendix*, Fig. S3). The difference is likely due to *gag* expression because it is the dominant product of proviruses missing internal genes (Fig. 4 *J* and *K*). As previously suggested (31), the disappearance of defectives that can produce *gag* protein could be mediated by cellular immunity.

Clonal Dynamics in the Intact Proviral Reservoir. To examine clonal dynamics in the intact proviral reservoir, we analyzed the intact HIV-1 proviral sequences from all participants across all time points. Overall, we obtained 152 intact proviral sequences from

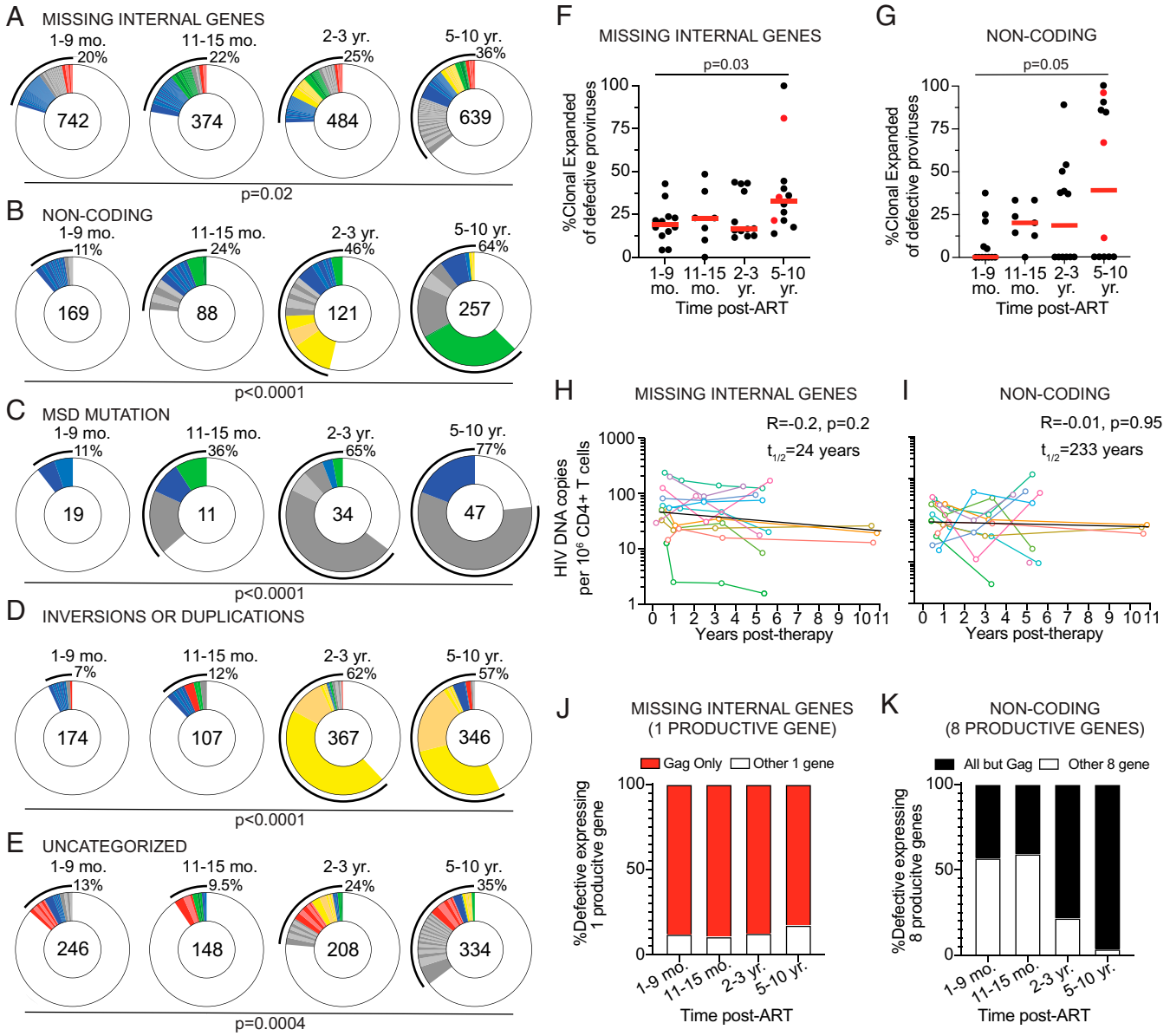

Fig. 4. Longitudinal analysis of clonal dynamics of distinct types of defective proviruses. (*A*–*E*) Pie charts showing persisting or expanded clones of defective proviruses categorized as (*A*) missing internal genes, (*B*) noncoding, (*C*) MSD mutation, (*D*) inversion or duplications, and (*E*) uncategorized. Time point assayed is denoted above the circles. Number inside each circle indicates the number of intact proviruses analyzed for each time point. Color slices indicate the following: Gray, clonally expanded defective proviruses unique to the time point assays; blue, persisting clones of defective provirus, first detected at 1 to 9 mo post-ART; green, persisting clones first detected at 11 to 15 mo post-ART; yellow, persisting clones first detected 2 to 3 y post-ART; red, persisting clones detected at all assayed time points; white, defective provirus sequences isolated only once. Black outline indicates the frequency of clonally expanded defective proviruses detected at each time point. (*F* and *G*) Frequency of clonally expanded defective proviruses per individual over time, when looking at defectives categorized as (*F*) missing internal genes and (*G*) noncoding. Participants sampled at 10 y post-ART are highlighted in red. Red horizontal line indicates median frequency. (*H* and *I*) Graph showing the number of defective HIV-1 proviruses per million $CD4^+$ T cells detected over time, when defectives are categorized as (*H*) missing internal (half-life = 24 y, $R = -0.2$, $P = 0.2$) and (*I*) noncoding (half-life = 233 y, $R = -0.01$, $P = 0.95$). Each colored line represents one individual participant. (*J*) Graphs showing the frequency of missing internal gene defective proviruses expressing one product where the ORF expressed is *gag*. (*K*) Graph showing the frequency of noncoding defective proviruses, which are sequence predicted to express eight functional proviral genes, where the ORFs expressed are all annotated genes excluding *gag*. A Kruskal–Wallis test with subsequent Dunn's multiple comparisons was used to analyze data where appropriate. The exponential decay half-life was calculated by using log10 IUPM values and a linear regression model.

samples collected 1 to 9 mo post-ART, 75 sequences from 11 to 15 mo post-ART, 67 sequences from 2 to 3 y, and 29 sequences from 5 to 10 y post-ART (Fig. 5*A*). The decrease in the number of intact sequences detected from samples obtained at the 5- to 10-y time point is consistent with the decay kinetics of the intact latent reservoir (Fig. 1*C*).

Expanded clones of $CD4^+$ T cells containing intact integrated proviruses were present in all but two of the individuals examined (Fig. 5*A*). As expected from prior work (15, 23, 34–36) the intact proviral reservoir was dynamic, with appearance and disappearance of some and persistence of other clones over time (Fig. 5*A*). When considered together the relative proportion of expanded clones in the latent reservoir increased with time from 15% at the first time point to 41% at the 5- to 10-y time point ($P < 0.0001$, Fig. 5*B*). Thus, clonality increases in parallel with the decrease in size of the intact

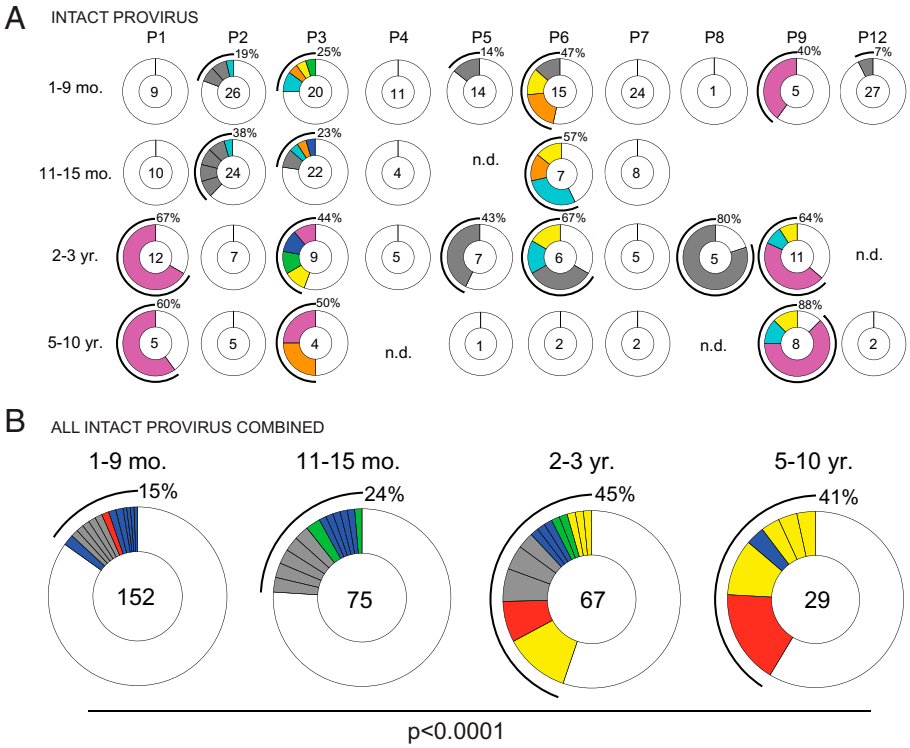

Fig. 5. Clonal dynamics of the intact provirus reservoir over time. (*A*) Pie charts of intact proviruses detected in each individual participant over time. Participant name is denoted above the circles. Number inside each circle indicates the number of intact proviruses analyzed for each time point, specified to the *Left* of the circles. Color slices indicate persisting intact clones found at multiple time points, while gray slices are clonal expansions unique to the time point. Remaining white slices represent sequences isolated only once. Black outlines indicate the frequency of clonally expanded intact proviruses detected in each participant. n.d., not detected, no intact proviruses were detected from this sample. (*B*) Pie chart showing clonality of all intact proviruses detected in all individuals, combined, over time. Time point assays are denoted above the circles. Number inside each circle indicates the number of intact proviruses analyzed for each time point. Color slices indicate the following: Gray, clonally expanded intact proviruses unique to the time point assays; blue, persisting clones of intact provirus, first detected at 1 to 9 mo post-ART; green, persisting clones first detected at 11 to 15 mo post-ART; yellow, persisting clones first detected 2 to 3 y post-ART; red, persisting clones detected at all assayed time points; and white, intact provirus sequences isolated only once. Black line indicates the frequency of clonally expanded intact proviruses detected at each time point. A Fisher's exact test with subsequent Bonferroni correction was used to compare clonality of intact proviruses between time points.

reservoir and as a result the complexity of the intact reservoir decreases with time. In conclusion, although the intact and defective reservoirs have different half-lives, they are both maintained in part by clonal expansion of $CD4^+$ T cells.

Discussion

We performed a longitudinal analysis of the HIV-1 latent reservoir using Q4PCR to gain insights into the longevity and sequence evolution of the intact and defective reservoir over time. Although Q4PCR differs from other methods in a number of ways, the results were similar to QVOA and IPDA in the following respects: 1) The decay rate of the intact reservoir measured by Q4PCR (4.9 y) was comparable to that measured by VOA (mean: 4.4 y, range: 2.3 to 8.1 y) and IPDA (4 to 18.5 y) (9–11, 15, 17, 21); 2) defective proviral decay was significantly slower than the intact reservoir (15, 17, 31); 3) there was no evidence for proviral replication in the latent reservoir (26, 37); 4) defective proviruses that have ORFs and can encode HIV-1 proteins decay faster than those that do not (27, 28, 31); and 5) increased defective proviral clonality over time (15, 16, 38). In addition, the analysis revealed that the clonality of the intact latent reservoir increases and its diversity decreases over time.

Differences between Q4PCR and other methods could be due to any number of pitfalls associated with each of the assays. For example, VOA is labor intensive, variable from assay to assay (10), and ultimately underestimates the latent reservoir (22, 23). IPDA is a sensitive, high-throughput assay (24), but is limited to detecting proviruses that are sequence compatible with its two packaging signal (PS) and *env* oligonucleotide probe sets (18, 39). In addition, IPDA lacks a sequence verification step to determine whether a provirus is truly intact or defective (20). The precision of the IPDA PS and *env* probes to identify genetically intact proviruses was estimated to be 70% in the original report (24). Similarly, a recently published direct comparison of IPDA and Q4PCR reported that of 625 probe-positive samples identified by the PS and *env* probe combination, 305 (49%) were defective (20). This discrepancy is variable between individuals, with some showing far larger or smaller differences than others (20). Nevertheless, the relative contribution of the more stable defective reservoir to the calculated IPDA half-life of the intact reservoir would be expected to increase over time as the intact reservoir decays. This potential problem may account for the increasing half-life of the intact reservoir measured by IPDA after 7 y and decrease the sensitivity of IPDA to changes in the intact reservoir over time (17).

Q4PCR is far more labor intensive than IPDA and not suited for the analysis of large clinical trials. In addition, the long-range PCR step in Q4PCR is less efficient that the short-range PCR used in IPDA, making Q4PCR a less sensitive assay that likely underestimates the size of the latent reservoir. We also limited our analysis to PBMCs isolated from blood samples. Although, blood appears to be representative of the overall latent reservoir (26, 40), it remains possible that we are missing information on reservoirs that reside solely in tissues. Despite these caveats, longitudinal studies show that half-life of the intact reservoir obtained by Q4PCR, VOA, and IPDA is remarkably similar.

Sequence analysis of the defective reservoir by near full-length sequencing on four individuals revealed that defective proviruses that encode proteins have a shorter half-life than those that cannot (31). These protein products could trigger immune responses leading to inflammation or T cell exhaustion (27, 28). In addition, peptide products of these proteins presented on human leukocyte antigen (HLA) may induce negative selection by cytotoxic $CD8^+$ T lymphocytes (28). In contrast, translationally silent proviruses might evade the immune system and thereby remain more stable over time (31). Thus, Q4PCR analysis confirms and extends prior observations, including the finding that the clonality of the defective reservoir increases over time (15, 16, 38).

It has been more challenging to examine the longitudinal clonal dynamics of the intact reservoir than the defective reservoir because of the relative difficulty in obtaining intact sequences at later time points. Q4PCR data reveal that clonality of the intact reservoir increases with time and as a result the reservoir becomes less diverse. The parallel increase in the proportion of proviral clones in the intact and defective reservoirs is consistent with the idea that clonal expansion of $CD4^+$ T cell harboring latent proviruses is driven in part by antigenic stimulation (32, 33).

In conclusion, Q4PCR is a low-throughput method that is complementary to VOA and IPDA. This method may be especially valuable for analysis of small studies that aim to document how clinical interventions may alter latent reservoir dynamics.

Materials and Methods

Study Participants. This study was approved by the institutional review board of the National Institute of Allergy and Infectious Diseases, NIH, and registered on ClinicalTrials.gov, NCT00039689. All study participants provided written informed consent, and leukapheresis products were collected in accordance with the protocol.

For two participants (P10 and P11), no intact proviruses were detected at the first time point (1 to 9 mo post-ART) after assaying genomic DNA (gDNA) from at least 0.5 M (million) $CD4^+$ T cells. For these participants, an average of 0.5 M $CD4^+$ T cells (range: 0.15 M to 1 M) were assayed per time point. For all remaining participants ($n = 10$), an average of 1.4 M (range: 0.5 M to 3.5 M) $CD4^+$ T cells were assayed per time point, resulting in a limit of detection of two intact proviruses per million $CD4^+$ T cells.

$CD4^+$ T Cell Isolation and Genomic DNA Extraction. Total $CD4^+$ T cells were isolated from cryopreserved PBMCs using magnetic labeling and negative selection using a $CD4^+$ T cell isolation kit (Miletnyi, #130-096-533). gDNA was immediately isolated from 3 to 37 million purified $C4^+$ T cells using phenol/chloroform (41). Briefly, $CD4^+$ T cells were lysed in Proteinase K buffer (100 mM Tris, pH 8, 0.2% sodium dodecyl sulphate [SDS], 200 mM NaCl, and 5 mM ethylenediaminetetraacetic acid [EDTA]) containing 20 mg/mL Proteinase K (Invitrogen, #AM2548), incubated overnight at 55 °C followed by gDNA extraction with phenol/chloroform/isoamyl alcohol extraction and ethanol participation, including an RNase A step. A Qubit 3.0 fluorometer and a Qubit dsDNA BR (broad-range) assay kit (Thermo Fisher Scientific, #Q32853) was used to measure DNA concentrations.

Q4PCR. Q4PCR was performed as previous described (25, 32). Briefly, an outer PCR (NFL1) was performed on gDNA at a single-copy dilution, which was determined by a gag-limiting dilution assay. Outer PCR primers include BLOuterF (5′-AAATCTCTAGCAGTGGCGCCCGAACAG -3′) and BLOuterR (5′-TGAGGGATCTCTAGTTACCAGAGTC-3′) (22). A total of 1 μL of NFL1 PCR product was used as a template for Q4PCR reaction, using a combination of four probes that target conserved regions of the HIV genome. Each probe set consisted of a forward and reverse primer pair, as well as a fluorescently labeled internal hydrolysis primer/probe, as previously described (25): PS forward (5′-TCTCTCGACGCAGGACTC-3′), PS reverse (5′-TCTAGCCTCCGCTAGTCAAA-3′), PS probe (5′-/Cy5/TTTGGCGTA/TAO/CTCACCAGTCGCC-3′/IAbRQSp, Integrated DNA Technologies), ENV forward (5′-AGTGGTGCAGAGAGAAAAAAGAGC-3′); ENV reverse (5′-GTCTGGCCTGTACCGTCAGC-3′), ENV probe (5′-/VIC/CCTTGGGTTCTTGGGA-3′/MGB, Thermo Fisher Scientific), Gag forward (5′-ATGTTTTCAGCATTATCAGAAGGA-3′), Gag reverse (5′-TGCTTGATGTCCCCCCACT-3′), Gag probe (5′-/6-FAM/CCACCCCAC/ZEN/AAGATTTAAACACCATGCTAA-3′/IABkFQ, Integrated DNA Technologies), and Pol forward (5′-GCACTTTAAATTTTCCCATTAGTCCTA-3′), Pol reverse (5′-CAAATTTCTACTAATGCTTTTATTTTTTC-3′), Pol probe (5′-/NED/AAGCCAGGAATGGATGGCC-3′/MGB, Thermo Fisher Scientific). Thermostable DNA polymerase was made in-house by transforming *Escherichia coli* BL21 (DE3) pLys S cells with pOpen Taw plasmid (Gene and Cell Technologies) and inducing the culture overnight with 1 mM isopropyl beta-D-1-thiogalactopyranoside (IPTG). Cell pellets were lysed with lysozyme and the enzyme was extracted, as previously published (42), followed by purification using a heparin Sepharose 6HR column. After dialysis, Taq polymerase was mixed with anti-Taq tp7 antibodies (43) at a 1:10 ratio. Each Q4PCR was performed in a 10-μL final reaction volume containing 0.1 μL of polymerase in DNA polymerase PCR buffer supplemented with $MgCl_2$ (Invitrogen 18067017), dNTP, Rox (Invitrogen #12223012), and containing the following primer and probe concentrations: PS forward and reverse primers at 21.6 μM with 6 μM of PS internal probe, *env* forward and reverse primers at 5.04 μM with 1.4 μM of *env* internal probe, *gag* forward and reverse primers at 2.7 μM with 0.75 μM *gag* internal problem, and lastly *pol* forward and reverse primers at 6.75 μM with 1.875 μM of *pol* internal probe. qPCR conditions were 94 °C for 10 min, 40 cycles of 94 °C for 15 s, and 60 °C for 60 s. All qPCR reactions were performed in a 384-well plate format using the Applied Biosystem QuantStudio 6 or 7 Flex Real-Time PCR system. We use QuantStudio Real-Time PCR software version 2.2 (Thermo Fisher Scientific) for data analysis. The same baseline correction (start cycle 3; end cycle 10) and normalized reporter signal (ΔRn) threshold (0.2) was set manually for all targets/probes. Fluorescent signal above the threshold was used to determine the threshold cycle. Samples with a threshold cycle value between 10 and 30 of any probe or probe combination were identified. Samples showing reactivity with two or more of the four qPCR probes were selected for a nested PCR (NFL2). The NFL2 was performed using 1 μL of the NFL1 product as a template. Reactions were formed in 20 μL reaction volume using Platinum Taq High Fidelity polymerase (Invitrogen, #11-304-029) and PCR primers 275F (5′-ACAGGGACCTGAAAGCGAAAG-3′) and 280R (5′-CTAGTTACCAGAGTCACACAACAGACG-3′) (22) at a concentration of 800 nM.

Library Preparation and Sequencing. All nested PCR products from NFL2 were subjected to library preparation, as previously described (32). The Qubit 3.0 Flourometer and Qubit dsDNA BR assay kit (Thermo Fisher Scientific, #Q32853) were used to measure DNA concentrations. Samples were diluted to a concentration of 10 to 20 ng/μL. Tagmentation reactions were performed using 1 μL of diluted DNA, 0.25 μL Nextera TDE1 Tagment DNA enzyme, and 1.25 μL TD Tagment DNA buffer (Illumina, #20034198). Tagmented DNA was ligated to a unique i5/i7 barcoded primer combination using the Illumina Nextera XT Index kit v2 and KAPA HiFi HotStart ReadyMix (Roche, #07958927001), and then purified using AmPure Beads XP (Beckman Coulter, #A63881). A total of 384 purified samples were pooled into one library and then subjected to paired-end sequencing using Illumina MiSeq Nano 300 V2 cycle kits (Illumina, #MS-103-1001) at a concentration of 12 pM.

HIV-1 Sequence Assembly and Annotation. HIV-1 sequence assembly was performed as previously described (32), using our in-house pipeline (Defective and Intact HIV Genome Assembler), which can reconstruct thousands of HIV genomes within hours via assembly of raw sequencing reads into annotated HIV genomes. The steps executed by the pipeline are described briefly as follows: First, we removed PCR amplification and performed error correction using clumpify.sh from BBTools and package v38.72 (https://sourceforge.net/projects/bbmap/). A quality-control check was performed with Trim Galore package v0.6.4 (https://github.com/FelixKrueger/TrimGalore) to trim Illumina adapters and low-quality bases. We also used bbduk.sh from the BBTools package to remove possible contaminant reads using HIV genome sequences, obtained from the Los Alamos HIV database, as a positive control. After the overlapping reads are merged by BBMerge, we use a k-mer–based assembler, SPAdes v3.13.1, to reconstruct the HIV-1 sequences. The longest assembled contig is aligned using BLAST with HXB2 to set it in the forward orientation. Some of the contig's nucleotides are potentially modified by aligning the high-quality HIV-derived reads. The modified contig is then annotated according to its alignment with HXB2 using BLAST. Sequences with double peaks, i.e., regions indicating the presence of two or more viruses in the sample (cutoff consensus identity for any residue <70%), or samples with a limited number of reads (empty wells ≤500 sequencing reads) were omitted from downstream analyses. In the end, sequences were classified as intact or defective, wherein intact sequences were determined by presence of 3′ long terminal repeat (LTR) and 5′ packaging signal (PSI) and lacking any fatal defects in the ORF of the nine genes annotated by sequence analysis. Defective sequences were subdivided into more specific classifications according to their sequence structure: MSD mutation, noncoding, missing internal genes, uncategorized, or inversions/duplications.

Clonal Analysis. Clones were defined by aligning sequences of each classification (intact, MSD mutation, noncoding, missing internal genes, uncategorized, or inversions/duplications) to HXB2 and calculating their Hamming distance. Sequences having a maximum of three differences between the first nucleotide of gag and last nucleotide of nef (reference: HXB2) were considered members of the same clone if found more than once across sequences from a single or multiple time points retrieved from each participant, based on the full amplification of a 9 KB template (99.667% identical).

Single-Genome Analysis (SGA) of Plasma Virus *env* Genes. Full-length *env* genes was amplified by nested PCR from plasma-derived viral cDNA, as previously described (44–46). Briefly, HIV-1 RNA was extracted from plasma samples using QIAamp Viral RNA Mini kit (Qiagen, #52904), followed by first-strand cDNA synthesis using SuperScript III reverse transcriptase (Invitrogen, #18-080-044). A total of 1 μL of cDNA diluted to single-genome per reaction was used as a template for first round *env* PCR, in a final reaction volume of 20 μL using Platinum Taq High Fidelity polymerase (Invitrogen, #11-304-029) and PCR primers envB5out (5′-TAGAGCCCTGGAAGCATCCAGGAAG-3′) and envB3out (5′-TTGCTACTTGTGATTGCTCCATGT-3′) at a concentration of 200 nM. Subsequently, a nested PCR using 1 μL of the first round PCR product in a similar PCR using PCR primers envB5in (5′-CACCTTAGGCATCTCCTATGGCAGGAAGAAG-3′) and envB3in (5′-GTCTCGAGATACTGCTCCCACCC-3′). PCR conditions were 94 °C for 2 min, 35 cycles of 94 °C for 15 s, 55 °C for 30 s, 68 °C for 4 min, followed by a single cycle of 68 °C for 15 min. Positive wells were determined using E-Gels (Invitrogen, #720841) and sequenced using Illumina MiSeq Nano 300 V2 cycle kits (Illumina, #MS-103-1001) and analyzed as described above. We assume all RNA recovered from plasma is reflective of intact viral particles. Sequence alignments, phylogenetic trees, and calculation of patristic distance to measure both the genetic distance and topology of the phylogenetic trees were performed by using Geneious Pro software, version 2020.0.3 and RAxML 8.2.11.

Statistical Analysis. Statistical analyses were performed using GraphPad Prism 9.

Data Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank all study participants who devoted time to our research, all members of the M.C.N. laboratory for helpful discussions, and Maša Jankovic for laboratory support. This work was supported by the Bill and Melinda Gates Foundation (Collaboration for AIDS Vaccine Discovery Grants OPOPP1092074, OPP1124068, and OPP1168933); the NIH (Grants 1UM1 AI100663 and R01AI129795) to M.C.N.; the Einstein–Rockefeller–CUNY Center for AIDS Research (Grant 1P30AI124414-01A1); REACH-HIV Delaney (Grant UM1 AI164565 to M.C.); and the Robertson Fund. T.-W.C. and S.M. were supported in part by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, NIH. C.G. was supported by the Robert S. Wennett Post-Doctoral Fellowship, and in part by the National Center for Advanced Translational Sciences (National Institutes of Health Clinical and Translational Science Award program, Grant UL1 TR001866), and by the Shapiro-Silverberg Fund for the Advancement of Translational Research. M.C.N. is a Howard Hughes Medical Institute investigator.

1. UNAIDS, *UNAIDS Data 2020.* https://www.unaids.org/en/resources/documents/2020/unaids-data (2020). Accessed 30 December 2021.
2. WHO, *World Health Organization: HIV/AIDS Key Facts.* https://www.who.int/en/news-room/fact-sheets/detail/hiv-aids (2020). Accessed 30 December 2021.
3. S. G. Deeks, J. Overbaugh, A. Phillips, S. Buchbinder, HIV infection. *Nat. Rev. Dis. Primers* **1**, 15035 (2015).
4. T. W. Chun *et al.*, Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* **387**, 183–188 (1997).
5. R. T. Davey Jr *et al.*, HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15109–15114 (1999).
6. J. K. Wong *et al.*, Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* **278**, 1291–1295 (1997).
7. T. W. Chun, S. Moir, A. S. Fauci, HIV reservoirs as obstacles and opportunities for an HIV cure. *Nat. Immunol.* **16**, 584–589 (2015).
8. M. J. Churchill, S. G. Deeks, D. M. Margolis, R. F. Siliciano, R. Swanstrom, HIV reservoirs: What, where and how to target them. *Nat. Rev. Microbiol.* **14**, 55–60 (2016).
9. N. Bachmann *et al.*, Swiss HIV Cohort Study, Determinants of HIV-1 reservoir size and long-term dynamics during suppressive ART. *Nat. Commun.* **10**, 1–11 (2019).
10. A. M. Crooks *et al.*, Precise quantitation of the latent HIV-1 reservoir: Implications for eradication strategies. *J. Infect. Dis.* **212**, 1361–1365 (2015).
11. J. D. Siliciano *et al.*, Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat. Med.* **9**, 727–728 (2003).
12. L. B. Cohn, N. Chomont, S. G. Deeks, The biology of the HIV-1 latent reservoir and implications for cure strategies. *Cell Host Microbe* **27**, 519–530 (2020).
13. K. M. Bruner *et al.*, Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat. Med.* **22**, 1043–1049 (2016).
14. S. Eriksson *et al.*, Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathog.* **9**, e1003174 (2013).
15. A. A. Antar *et al.*, Longitudinal study reveals HIV-1-infected CD4+ T cell dynamics during long-term antiretroviral therapy. *J. Clin. Invest.* **130**, 3543–3559 (2020).
16. L. B. Cohn *et al.*, HIV-1 integration landscape during latent and active infection. *Cell* **160**, 420–432 (2015).
17. M. J. Peluso *et al.*, Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy. *JCI Insight* **5**, e132997 (2020).
18. F. R. Simonetti *et al.*, Intact proviral DNA assay analysis of large cohorts of people with HIV provides a benchmark for the frequency and composition of persistent proviral DNA. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 18692–18700 (2020).
19. R. T. Gandhi *et al.*, AIDS Clinical Trials Group A5321 Team, Selective decay of intact HIV-1 proviral DNA on antiretroviral therapy. *J. Infect. Dis.* **223**, 225–233 (2021).
20. C. Gaebler *et al.*, Sequence evaluation and comparative analysis of novel assays for intact proviral HIV-1 DNA. *J. Virol.* **95**, e01986-20 (2021).
21. D. Finzi *et al.*, Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat. Med.* **5**, 512–517 (1999).
22. Y. C. Ho *et al.*, Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155**, 540–551 (2013).
23. N. N. Hosmane *et al.*, Proliferation of latently infected $CD4^+$ T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J. Exp. Med.* **214**, 959–972 (2017).
24. K. M. Bruner *et al.*, A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature* **566**, 120–125 (2019).
25. C. Gaebler *et al.*, Combination of quadruplex qPCR and next-generation sequencing for qualitative and quantitative analysis of the HIV-1 latent reservoir. *J. Exp. Med.* **216**, 2253–2264 (2019).
26. W. R. McManus *et al.*, HIV-1 in lymph nodes is maintained by cellular proliferation during antiretroviral therapy. *J. Clin. Invest.* **129**, 4629–4642 (2019).
27. H. Imamichi *et al.*, Defective HIV-1 proviruses produce viral proteins. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 3704–3710 (2020).
28. R. A. Pollack *et al.*, Defective HIV-1 proviruses are expressed and can be recognized by cytotoxic T lymphocytes, which shape the proviral landscape. *Cell Host Microbe* **21**, 494–506.e4 (2017).
29. B. Cole *et al.*, In-depth single-cell analysis of translation-competent HIV-1 reservoirs identifies cellular sources of plasma viremia. *Nat. Commun.* **12**, 1–13 (2021).
30. G. Sannier *et al.*, Combined single-cell transcriptional, translational, and genomic profiling reveals HIV-1 reservoir diversity. *Cell Rep.* **36**, 109643 (2021).
31. M. R. Pinzone *et al.*, Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by clonal expansion. *Nat. Commun.* **10**, 1–12 (2019).
32. P. Mendoza *et al.*, Antigen-responsive CD4+ T cell clones contribute to the HIV-1 latent reservoir. *J. Exp. Med.* **217**, e20200051 (2020).
33. F. R. Simonetti *et al.*, Antigen-driven clonal selection shapes the persistence of HIV-1 infected CD4+ T cells in vivo. *J. Clin. Invest.* **131**, 145254 (2021).
34. J. K. Bui *et al.*, Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS Pathog.* **13**, e1006283 (2017).
35. J. C. Lorenzi *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.* **113**, E7908–E7916 (2016).
36. Z. Wang *et al.*, Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2575–E2584 (2018).
37. M. R. Abrahams *et al.*, The replication-competent HIV-1 latent reservoir is primarily established near the time of therapy initiation. *Sci. Transl. Med.* **11**, eaaw5589 (2019).
38. T. A. Wagner *et al.*, An increasing proportion of monotypic HIV-1 DNA sequences during antiretroviral treatment suggests proliferation of HIV-infected cells. *J. Virol.* **87**, 1770–1778 (2013).
39. N. N. Kinloch *et al.*, HIV-1 diversity considerations in the application of the Intact Proviral DNA Assay (IPDA). *Nat. Commun.* **12**, 165 (2021).
40. A. Chaillon *et al.*, HIV persists throughout deep tissues with repopulation from multiple anatomical sources. *J. Clin. Invest.* **130**, 1699–1712 (2020).
41. I. A. Klein *et al.*, Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* **147**, 95–106 (2011).
42. F. G. Pluthero, Rapid purification of high-activity Taq DNA polymerase. *Nucleic Acids Res.* **21**, 4850–4851 (1993).
43. J. L. Daiss, E. R. Scalice, D. J. Sharkey, Topographical characterization of the DNA polymerase from Thermus aquaticus. Defining groups of inhibitor mAbs by epitope mapping and functional analysis using surface plasmon resonance. *J. Immunol. Methods* **183**, 15–26 (1995).
44. M. Caskey *et al.*, Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nat. Med.* **23**, 185–191 (2017).
45. J. F. Salazar-Gonzalez *et al.*, Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J. Virol.* **82**, 3952–3970 (2008).
46. J. F. Scheid *et al.*, HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption. *Nature* **535**, 556–560 (2016).