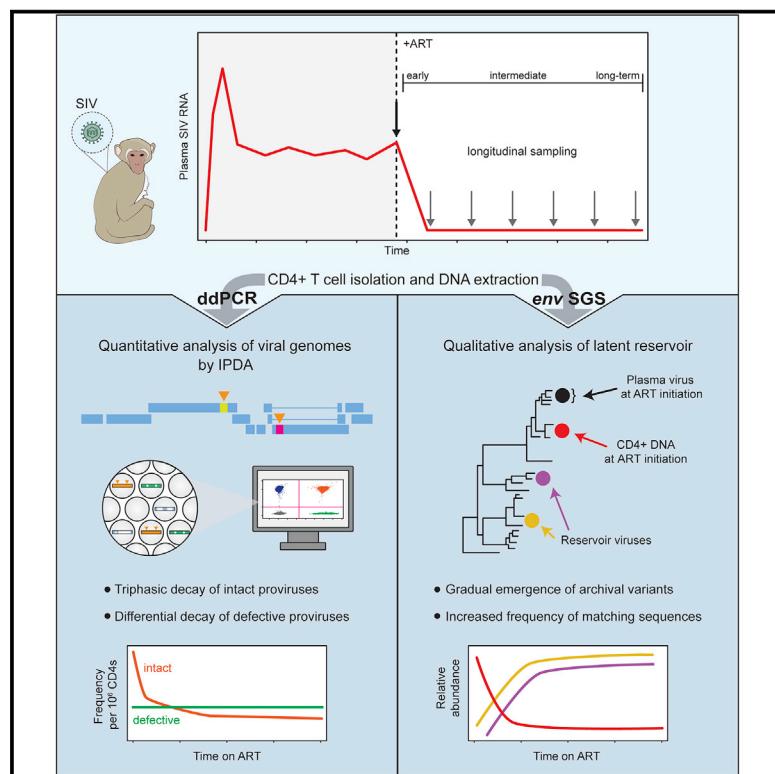


Antiretroviral therapy reveals triphasic decay of intact SIV genomes and persistence of ancestral variants

Graphical abstract



Authors

Emily J. Fray, Fengting Wu,
Francesco R. Simonetti, ...,
Dan H. Barouch, Janet D. Siliciano,
Robert F. Siliciano

Correspondence

efray1@jhmi.edu (E.J.F.),
rsiliciano@jhmi.edu (R.F.S.)

In brief

SIV-infected macaques are an important animal model for people living with HIV (PLWH). Fray et al. measured the decay of SIV-infected cells after treatment with antiretroviral therapy. They found that SIV-infected cells decay more rapidly than in PLWH and provide a benchmark for evaluating future cure strategies in this model.

Highlights

- Intact SIV genomes decay with triphasic kinetics upon treatment with ART
- Defective (hypermutated) proviruses decay with mono or biphasic kinetics
- Early after treatment, labile populations of recently infected cells decay
- Long-term ART reveals persistence of archival variants and infected cell proliferation



Article

Antiretroviral therapy reveals triphasic decay of intact SIV genomes and persistence of ancestral variants

Emily J. Fray,^{1,7,*} Fengting Wu,¹ Francesco R. Simonetti,¹ Carolin Zitzmann,² Narmada Sambaturu,² Carmen Molina-Paris,² Alexandra M. Bender,¹ Po-Ting Liu,³ John D. Ventura,³ Roger W. Wiseman,⁴ David H. O'Connor,⁴ Romas Gelezinius,⁵ Thomas Leitner,² Ruy M. Ribeiro,² Alan S. Perelson,² Dan H. Barouch,³ Janet D. Siliciano,¹ and Robert F. Siliciano^{1,6,*}

¹Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

²Los Alamos National Laboratory, Los Alamos, NM 87545, USA

³Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA

⁴Wisconsin National Primate Research Center, Madison, WI 53715, USA

⁵Gilead Sciences, Foster City, CA 94404, USA

⁶Howard Hughes Medical Institute, Baltimore, MD 21205, USA

⁷Lead contact

*Correspondence: efray1@jhmi.edu (E.J.F.), rsiliciano@jhmi.edu (R.F.S.)

<https://doi.org/10.1016/j.chom.2023.01.016>

SUMMARY

The decay kinetics of HIV-1-infected cells are critical to understand virus persistence. We evaluated the frequency of simian immunodeficiency virus (SIV)-infected cells for 4 years of antiretroviral therapy (ART). The intact proviral DNA assay (IPDA) and an assay for hypermutated proviruses revealed short- and long-term infected cell dynamics in macaques starting ART ~1 year after infection. Intact SIV genomes in circulating CD4⁺T cells showed triphasic decay with an initial phase slower than the decay of the plasma virus, a second phase faster than the second phase decay of intact HIV-1, and a stable third phase reached after 1.6–2.9 years. Hypermutated proviruses showed bi- or mono-phasic decay, reflecting different selective pressures. Viruses replicating at ART initiation had mutations conferring antibody escape. With time on ART, viruses with fewer mutations became more prominent, reflecting decay of variants replicating at ART initiation. Collectively, these findings confirm ART efficacy and indicate that cells enter the reservoir throughout untreated infection.

INTRODUCTION

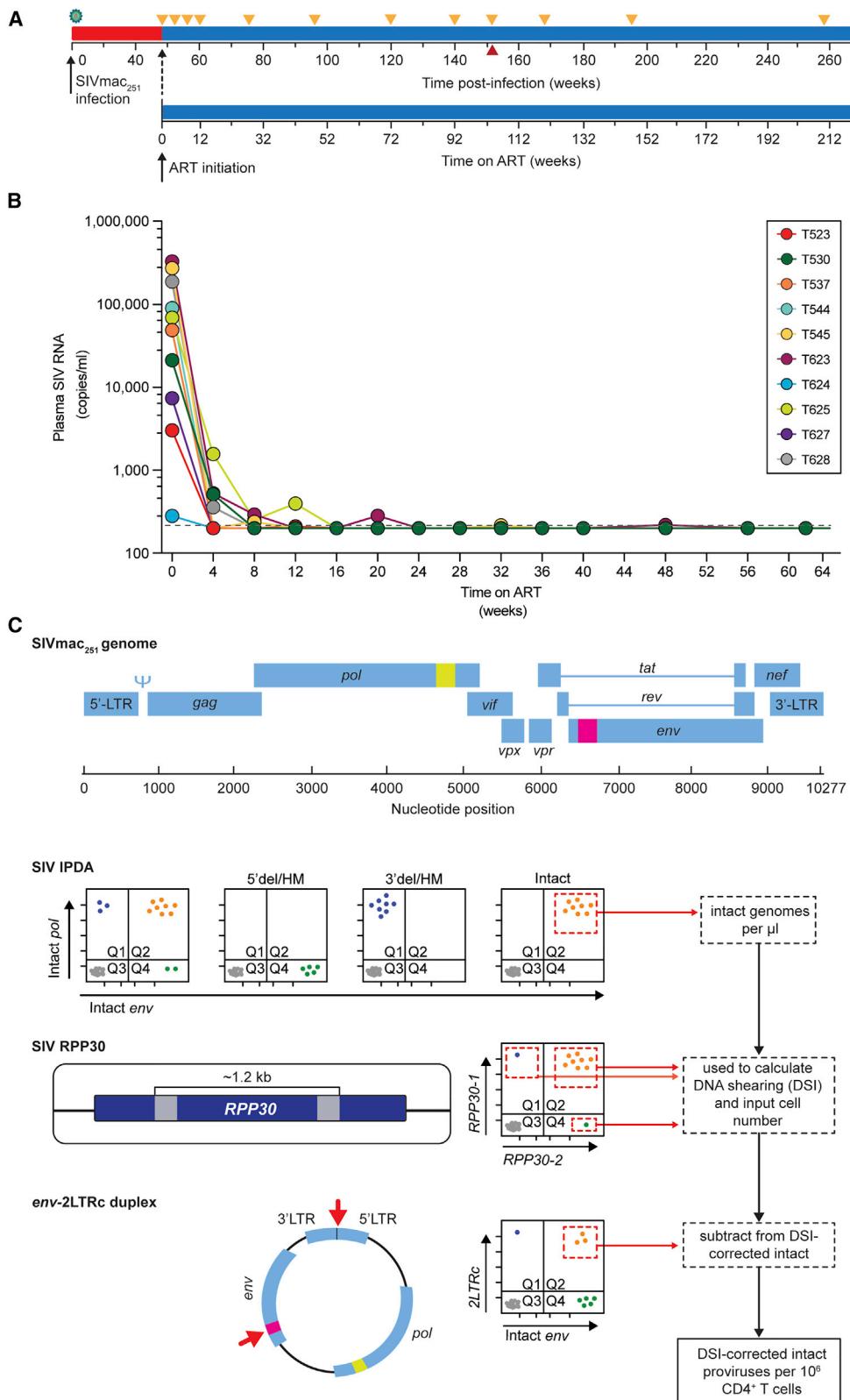
Antiretroviral therapy (ART) halts HIV-1 replication, resulting in a rapid biphasic drop in viremia to below the detection limit.¹ However, ART is not curative due to a pool of latently infected CD4⁺ T cells that harbor replication-competent transcriptionally silent proviruses.^{2–4} This small population of cells has a half-life of 3.7 years, guaranteeing lifetime HIV-1 persistence even with optimal ART.^{5,6} Regardless of the length of treatment, viremia rapidly rebounds if ART is stopped.^{7,8}

The decay kinetics of plasma virus and infected CD4⁺ T cells provide insights into the establishment and persistence of the latent reservoir.^{1,9,10} However, the measurement of infected cell populations is complicated because in people living with HIV-1 (PLWH) on ART, most infected cells harbor proviruses with fatal defects, including large deletions and APOBEC3G/F-mediated G-to-A hypermutation.^{11–19} To address this problem, we developed a digital droplet PCR (ddPCR) assay in which individual proviruses suspended in nanoliter-sized droplets are

interrogated with two strategically chosen amplicons that are more likely to detect intact viral genomes.^{20,21} This intact proviral DNA assay (IPDA) provides a more accurate measurement of the reservoir size by excluding most defective proviruses.

The IPDA was used to confirm the long-term stability of the reservoir^{21–24} and more recently, to measure the decay of intact and defective HIV-1 genomes during the first 2 years of ART.²⁵ Some intact genomes detected at ART initiation may be linear unintegrated viral DNA molecules in resting cells.^{26,27} However, these forms decay within days,^{28,29} and in activated cells, integration occurs within hours of completion of reverse transcription.³⁰ Thus, most viral genomes detected with the IPDA are integrated proviruses. White et al. used the IPDA to show that after ART initiation intact genomes decayed with biphasic kinetics characterized by a rapid initial phase ($t_{1/2} = 12.9$ days) and a slower second phase ($t_{1/2} = 19$ months) that was still faster than the slow reservoir decay observed in PLWH on long-term ART.^{5,6,24,25} In this same study, cells with defective proviruses showed monophasic decay, and 2LTR circles, which are also





(legend on next page)

defective for replication showed biphasic decay and long-term stability.^{25,31–34} These findings raise questions about the nature and location of the populations of infected cells that produce most of the plasma virus, the mechanisms by which they decay, and the relationship between these rapidly decaying populations and the latent reservoir.

These questions can be more easily addressed in macaques infected with simian immunodeficiency virus (SIV) because the sequence of the infecting virus and the timing of infection and ART can be controlled.^{35–38} As in PLWH, there is viremia, progressive CD4⁺ T cell depletion, chronic immune activation, decay of viremia upon ART initiation, immune control in some instances, and a latent reservoir established in resting CD4⁺ T cells.^{36,39–49} Measurement of the SIV latent reservoir is also complicated by defective viral genomes.^{20,50} The decay of SIV-infected cells in macaques on ART has not been fully described.

Here, we evaluated changes in the frequency of SIV-infected cells for 4 years following ART initiation. We used the IPDA and an assay for hypermutated proviruses to compare the decay rates of intact and defective proviruses. We used single-genome sequencing (SGS) of plasma virus and proviral DNA to examine qualitative changes in infected cell populations during ART. The sequencing analyses provided strong evidence for the efficacy of ART and insights into whether the latent reservoir is established predominantly at ART initiation or continuously throughout untreated infection.^{51–55} The changes in infected cell populations identified here should be considered in the design and evaluation of cure interventions.

RESULTS

Cohort characteristics

We studied Indian-origin rhesus macaques (cohort 18–02, n = 10) inoculated with SIVmac₂₅₁ by repetitive low-dose intrarectal challenge (Figures 1A and S1). After 48 weeks, animals were placed on a combination ART regimen of tenofovir disoproxil fumarate, emtricitabine, and dolutegravir (TDF/FTC/DTG).^{56,57} They were maintained on ART without additional interventions for >4 years (Figure 1A). Cohort details are included in the supplement and the STAR Methods section (Figure S1A).

Changes in plasma virus upon ART initiation

All animals were infected by SIVmac₂₅₁ after one or more challenges and showed peak viremia typical of SIV infection of rhesus macaques (geometric mean = 1.05×10^7 copies/mL) (Figure S1A).⁴⁰ ART rapidly suppressed viral replication as

demonstrated by an exponential decrease in plasma SIV RNA (Figures 1B and S1B). For 6 of 10 animals, plasma SIV RNA levels fell below the detection limit within 4 weeks (Figures 1B and S1B). 4 of 10 animals experienced infrequent blips during the first 48 weeks of ART, after which plasma SIV RNA remained undetectable (Figure 1B).

IPDA analysis of reservoir size and changes in intact proviruses on ART

According to classic viral dynamics models,^{1,9,10,58} the rapid drop in viremia following ART initiation (Figure 1B) reflects the loss of productively infected cells. However, this prediction is difficult to verify because cells that produce most of the plasma virus may not be in the circulation,^{25,59} and because many cells carry defective proviruses.^{20,21,11} Therefore, we used the SIV IPDA²⁰ to quantitate cells with intact proviruses. This assay excludes most deleted or hypermutated proviruses.^{20,21} Our analysis also included a duplex ddPCR assay for a cellular gene (RPP30) and a duplex ddPCR assay for 2LTR circles (Figure 1C).^{20,60}

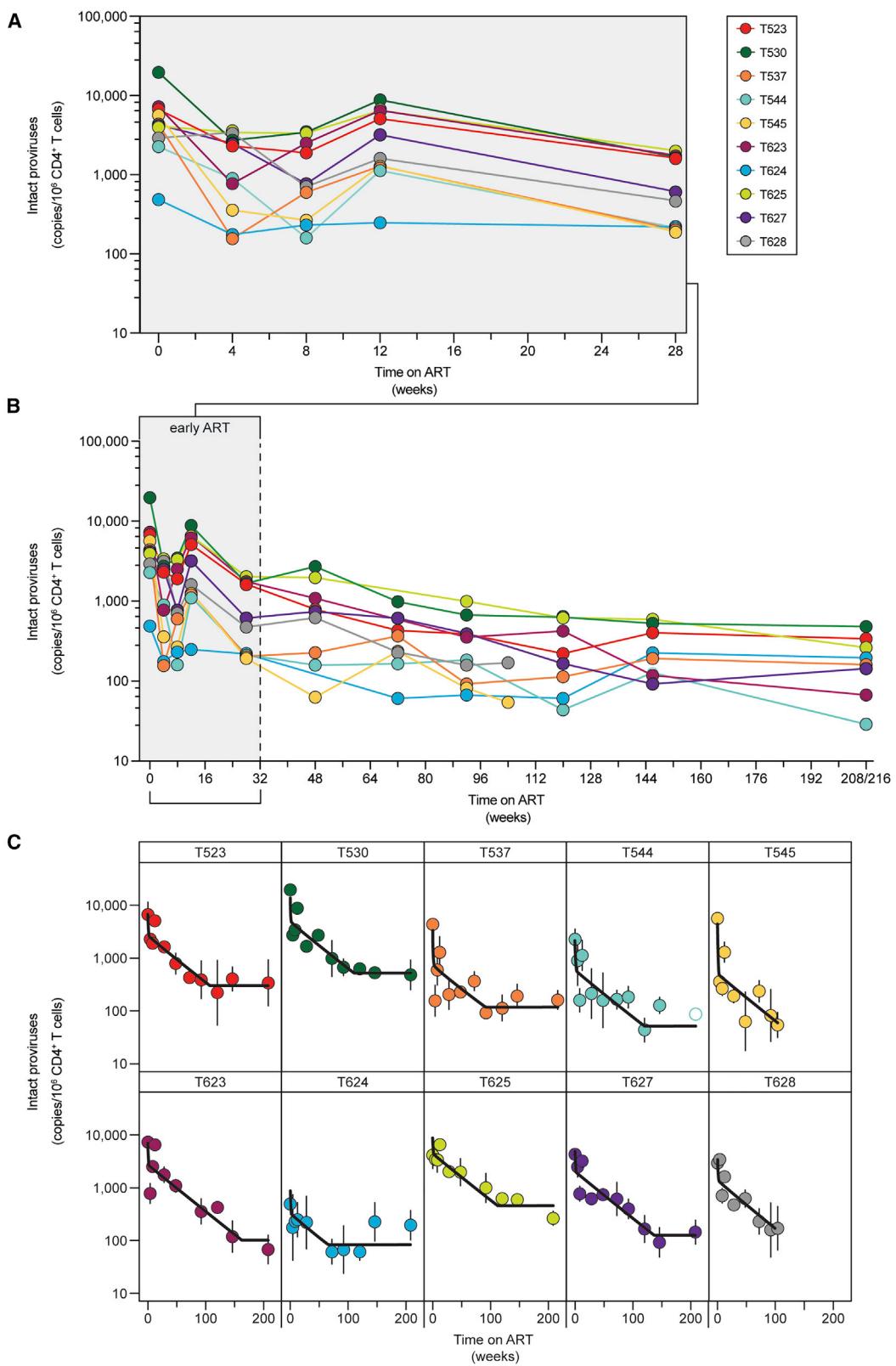
At ART initiation, intact proviral frequencies in circulating CD4⁺ T cells were similar to those in untreated PLWH (Figures 2A–2C, geometric mean frequency 3,642 copies per 10^6 CD4⁺ T cells, compared with 2,255 copies per 10^6 CD4⁺ T cells for HIV-1²⁵). We did not observe a relationship between viral load at week 0 and the frequency of intact proviruses at either week 0 or at a later time point (week 92) after which the more labile populations of infected cells have largely decayed (Figures S1B and S1C). We observed rapid initial decay of intact proviruses over the first 4 weeks of ART, a small but consistent increase around 12 weeks (Figure 2A), then slower decay for ~2 years until stabilization occurred (Figure 2B). We used non-linear mixed-effects modeling to determine the decay rate of intact proviruses. Changes in the frequency of cells harboring intact genomes were best described by a three-phase model, which is most evident in plots for individual macaques (Figure 2C; Table S1).

During the first 4 weeks of ART, intact SIV proviruses decayed with a half-life of 3.3 days (95% CI: 2.5–4.5 days), faster than the initial decay of HIV-1 proviruses in people initiating ART (12.9 days²⁵). Although fast compared with the subsequent decay of the remaining intact SIV proviruses, it is nevertheless slower than the decay of plasma SIV RNA over the same period (Figures S1D and S1E). Therefore, the very rapid initial decay of viremia may reflect the equally rapid death of infected cells located primarily in tissues, as seen for HIV-1.²⁵

Given the strong suppression of viremia upon ART initiation, the increase in intact proviruses at week 12 was unexpected.

Figure 1. Study design

- (A) Timeline of infection, initiation of ART, and sampling. Animals were infected by repetitive intrarectal challenge with SIVmac₂₅₁ and maintained untreated for 48 weeks after initial challenge (red) before initiation of ART (blue). Orange triangles indicate sampling times. The red triangle indicates the last time point before two animals (T545 and T628) were necropsied due to unrelated complications.
- (B) Plasma SIV RNA levels after ART initiation. The assay detection limit (200 copies/mL) is indicated by the dashed gray line.
- (C) Design of the SIV intact proviral DNA assay (IPDA).²⁰ Locations of two amplicons in the SIV genome are indicated in yellow (*pol*) and pink (*env*). Data from 3 individual assays are used to determine the frequency of intact proviruses per 10^6 CD4⁺ T cells. Events from quadrant 2 (Q2) of the IPDA measure the number of intact genomes. Proviruses with deletions or hypermutation affecting the *pol* amplicon (5'del/HM) appear in Q4 while proviruses with deletions or hypermutation affecting the *env* amplicon (3'del/HM) appear in Q1. Proviruses with defects affecting both amplicons appear in Q3 along with droplets lacking a provirus. The number of intact proviruses (Q2) is corrected for shearing using the DNA Shearing Index (DSI), and the input cell number is determined using ddPCR amplification of two similarly spaced amplicons in a cellular gene (RPP30) as described.^{20,21} 2LTR circles are quantitated in a separate ddPCR assay using two amplicons, one spanning the LTR-LTR junction⁶⁰ and one in *env* (red arrows).²⁰ Events in Q2 of the *env*-2LTRc duplex assay are subtracted from the number of intact proviruses to correct for 2LTR circles.



(legend on next page)

Single-genome sequencing described below did not reveal any unique properties of variants detected at week 12, such as an increase in divergence. This increase at approximately the same time post-treatment has also been observed in an independent cohort of SIV-infected animals (unpublished data, Fray) and may reflect a redistribution of infected cells from tissues to blood.^{61,62} In modeling the decay kinetics, we tested models that included (Table S1) or ignored the week 12 time point (Table S2). Whether the week 12 data were included or not, the best fit model was consistent with a triphasic decay pattern. All results discussed below are based on the full dataset (Table S1). It is unclear whether this increase in intact proviruses also occurs in animals that are started on ART after the relatively short periods of untreated infection characteristic of many cure studies.

The rapid initial decay of intact SIV proviruses was followed by a slower second phase ($t_{1/2} = 8.1$ months, 95% CI: 6.5–10.0 months). This decay is also faster than the second phase decay of intact HIV-1 proviruses ($t_{1/2} = 19$ months²⁵). The more rapid first and second phase decay of intact SIV genomes relative to HIV-1 has not been previously described. After a mean of 2.3 years (95% CI: 1.6–2.9 years), the slope changes to a stable third phase with no decay (Figure 2C; Tables S1 and S3). Most SIV cure experiments are conducted during the first 2 years of ART, when the more rapid second phase decay is still occurring. Thus, adequately powered control groups are needed to distinguish natural and intervention-induced decay.

Decay of 2LTR circles during ART

2LTR circles showed triphasic decay similar to that measured for intact proviruses (Figure S2; Tables S1 and S4). After 4 years of ART, 2LTR circles were rare. *env*⁺2LTR circles, which are subtracted from the Q2 IPDA result, were only detected in 2 of 8 animals at weeks 208–216 (Figure S2). The low frequency of *env*⁺2LTR circles after 4 years of ART could reflect the death of cells harboring the circles and/or cellular proliferation causing dilution of the circles, which are not copied during replication of host cells. Overall, total and *env*⁺ 2LTR circles showed decay kinetics that were very similar to those described for intact proviruses (Figure S2; Tables S1 and S3), as is the case with HIV-1.²⁵

An assay for hypermutated proviruses

In addition to detecting intact proviruses in Q2, the HIV-1 IPDA detects proviruses with hypermutation and/or deletions in Q1 and proviruses with 5' deletions in Q4.²¹ Because of differences in the distribution of defects, the SIV IPDA uses different amplicons,²⁰ and many of the defective proviruses end up in Q3 (Figure 1C) with the droplets that do not contain a provirus (estimated ~55% compared with ~3.8% for HIV based on *in silico* analyses of near-full-length sequencing data^{20,21}). For this reason, the SIV IPDA does not provide quantitation of defective proviruses. To determine whether decay of intact proviruses is

due to virus-dependent factors such as viral cytopathic effects and immune clearance or to virus-independent factors such as normal CD4⁺ T cell turnover, we developed a method to quantify one class of defective SIV genomes. APOBEC3G/F-mediated hypermutation introduces multiple missense and nonsense mutations in most open reading frames, likely precluding high-level viral gene expression.²⁰ We therefore developed a ddPCR assay specific for hypermutated SIV proviruses based on the same principles and amplicon positions as the SIV IPDA. We used fluorescently labeled *pol* and *env* probes with 2 G → A mutations at frequently hypermutated positions within each probe and unlabeled competitor probes with the non-mutated sequences (Figure S3). Details of the hypermutated proviral DNA assay (HPDA) are provided in the supplement (Figure S3).

Hypermutated proviruses were less frequent than intact proviruses (geometric mean frequency = 539 copies per 10^6 CD4⁺ T cells, compared with 3,642 intact copies per 10^6 CD4⁺ T cells, Figure 3A). Hypermutated proviruses were most abundant at the time of ART initiation (Figure 3A) and decayed with kinetics different from those of intact proviruses (Figure 3B) and variable between animals (Table S5). The best fit model for hypermutated proviruses was not uniform. There was biphasic decay for 5 animals and monophasic decay for the other 5 (Figure 3B; Table S5). For the animals with biphasic decay, hypermutated proviruses decayed initially with an average half-life of 42 days (95% CI: 31–55 days), followed by a stable second phase with no decay (Figure 3B; Tables S1 and S5). For the other 5 animals, we observed monophasic decay with an average half-life of 27.8 months (95% CI: 19.5–39.9 months) (Figure 3B; Tables S1 and S5). These kinetics are strikingly different from those described for cells harboring intact SIV proviruses. Overall, the decay of hypermutated proviruses is slower, such that by week 216 the average frequencies are similar (Figure 3A).

Analysis of infected cell decay using *Gag* PCR

Many SIV reservoir studies use a single amplicon in the well-conserved *gag* gene to quantify SIV DNA.^{63,64} We therefore analyzed viral decay using a ddPCR assay targeting *gag*.⁶⁵ This assay gave values similar to, or slightly higher than those measured by the IPDA (Figure 4A). This reflects the fact that the ratio of intact to defective proviruses is higher for HIV-1 than for SIV^{20,21,50} and the fact that some defective SIV proviruses have hypermutation or deletions affecting the *gag* amplicon and thus are also excluded by the *gag* PCR assay. We found a strong positive correlation between the *gag* and IPDA assays ($r = 0.8812$, $R^2 = 0.7765$, 95% CI = 0.8281–0.9186, $p < 0.0001$, $n = 100$; Pearson's test). Even when the *gag* data were corrected by the same *env*⁺2LTR correction applied to the IPDA values (Figure 4B; $r = 0.8854$, $R^2 = 0.7840$, 95% CI = 0.8341–0.9216, $p < 0.0001$, $n = 100$; Pearson's test), *gag* values were, in general, higher, likely reflecting the detection of some defective proviruses with the *gag* assay.

Figure 2. Decay of intact proviruses after ART initiation

(A and B) Frequency of intact proviruses during (A) the first 6 months of ART or (B) over the first 4 years of treatment. Datapoints represent the geometric mean from at least 3 replicates per animal. Open circles indicate that no positive events were detected by the IPDA; the y axis value indicates the limit of detection based on the number of cell equivalents assayed.

(C) Decay of intact proviruses in individual animals during ART. Data were fitted to a triphasic model (black line) as described in the STAR Methods. Vertical lines indicate the standard deviations.

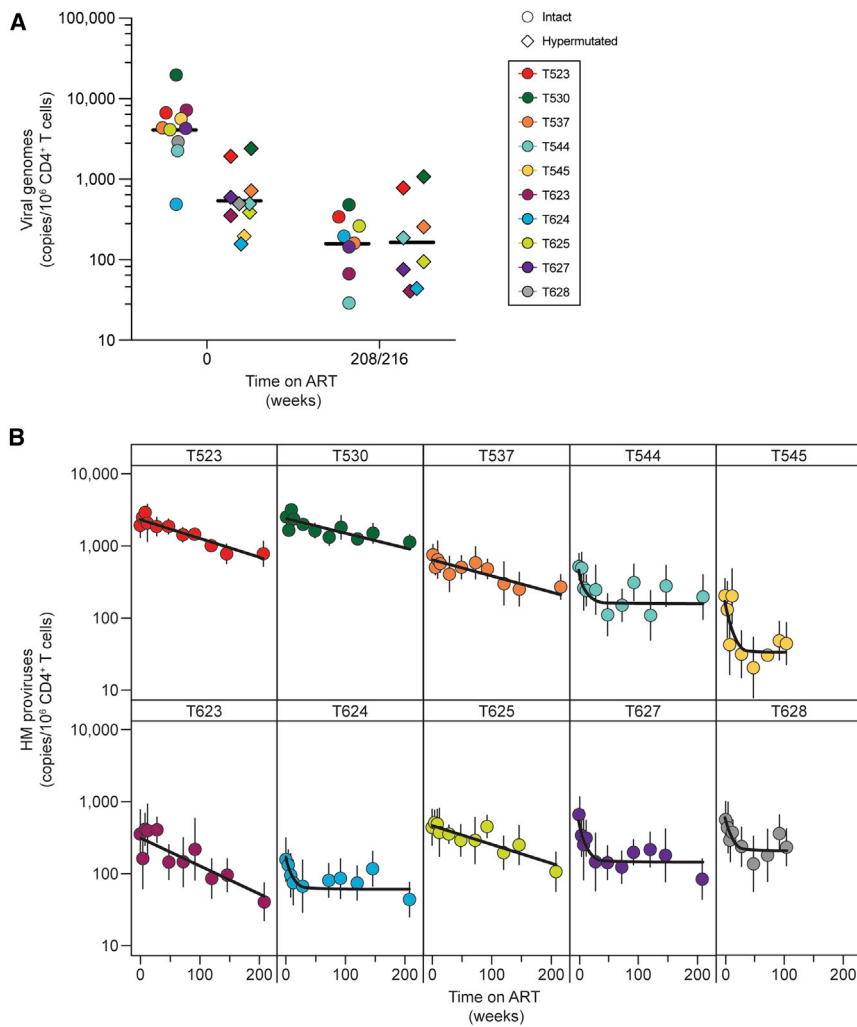


Figure 3. Decay of hypermutated proviruses after ART initiation

(A) Comparison of the frequency of intact and hypermutated proviruses prior to ART initiation ($t = 0$) and at the latest time point analyzed (week 208 for all animals except T537, in which case it is week 216). Each data point represents the geometric mean for the IPDA or HPDA. A total of 3 replicates for each animal are included for the IPDA, and 6 replicates for the HPDA. IPDA data are corrected for env^+2LTR circles and shearing. HPDA data are corrected for shearing. Black lines show the geometric means.

(B) Decay of hypermutated proviruses on ART. Data were fitted to biphasic or monophasic decay models (black line) as described in the STAR Methods section. Vertical lines represent the standard deviations.

defective proviruses, but overestimates the frequency of intact proviruses, especially after long-term ART.

Qualitative changes in infected cell populations on ART

To understand qualitative changes in the pool of SIV-infected cells following ART initiation, we collected a large number of env sequences by single-genome sequencing of DNA from peripheral blood CD4⁺ T cells at weeks 0–216 of ART (Figure 5A). After excluding sequences with deletions, frameshifts, or hypermutation, we were left with ~1,700 sequences that were non-defective in the env gene. We determined the average pairwise distance (APD) from the consensus sequence of the inoculum.⁵⁶ We found no increase in APD over ~4 years of treatment (Figure 5B; Table S7). Thus, despite the enormous diversification and divergence away from the stock sequence during untreated infection, there was no further evolution during ART. In fact, APD decreased with increasing time on ART (Figure 5C). The mechanism underlying this apparent “reverse evolution” is described below.

We also analyzed env sequences for evidence of proliferation of infected cells, as has been clearly demonstrated in PLWH.^{66–70} In macaques infected with SIVmac₂₅₁ for ~1.5 years and treated with ART for 60 weeks,²⁰ we previously found only 2 pairs of identical sequences among more than 500 independent env sequences. Another study used integration site analysis to recover matching integration sites from macaques treated during acute infection, although at low abundance.⁷¹ The extended follow-up here allowed us to observe the appearance of identical sequences over time. Among ~1,700 non-defective sequences from 9 animals, the fraction of identical sequences increased with time on ART (Figure 5C) as seen in PLWH.^{12,70} However, proviruses with identical sequences could contain polymorphisms elsewhere.^{72,73} Therefore, we extended our analysis to include 226 defective, hypermutated

As with intact proviruses, the decay of gag^+ DNA was best described by a triphasic model (Figure 4C; Tables S1 and S6). The average first phase half-life was 6.5 days (95% CI: 5.9–7.4 days), not significantly different from the initial decay of intact proviruses. However, at the later time points the half-lives became longer relative to intact proviruses (Table S6). We calculated an average second phase half-life of 13 months (95% CI: 11.8–14.1 months; Table S1). After a mean of 34.5 months (95% CI: 27.3–43.3 months) there was a third phase with essentially no decay (Tables S1, S6). These results are consistent with the fact that the gag assay is also capturing some defective proviruses, which generally have a longer half-life.^{24,25} Indeed, the gag probe detected hypermutated proviruses with 1 to 3 G→A mutations in the probe-binding site (Figures S4A–S4C). Consistent with this observation, the greatest differential between gag and IPDA measurements was found in animals with a higher frequency of hypermutated proviruses detected by the HPDA (animals T523, T530, and T537; Figures 4C, S4B, and S4D), although as time on ART increased, there was a greater discrepancy between the two assays in most animals. Therefore, the gag assay is a good approximation of the total abundance of viral DNA, including

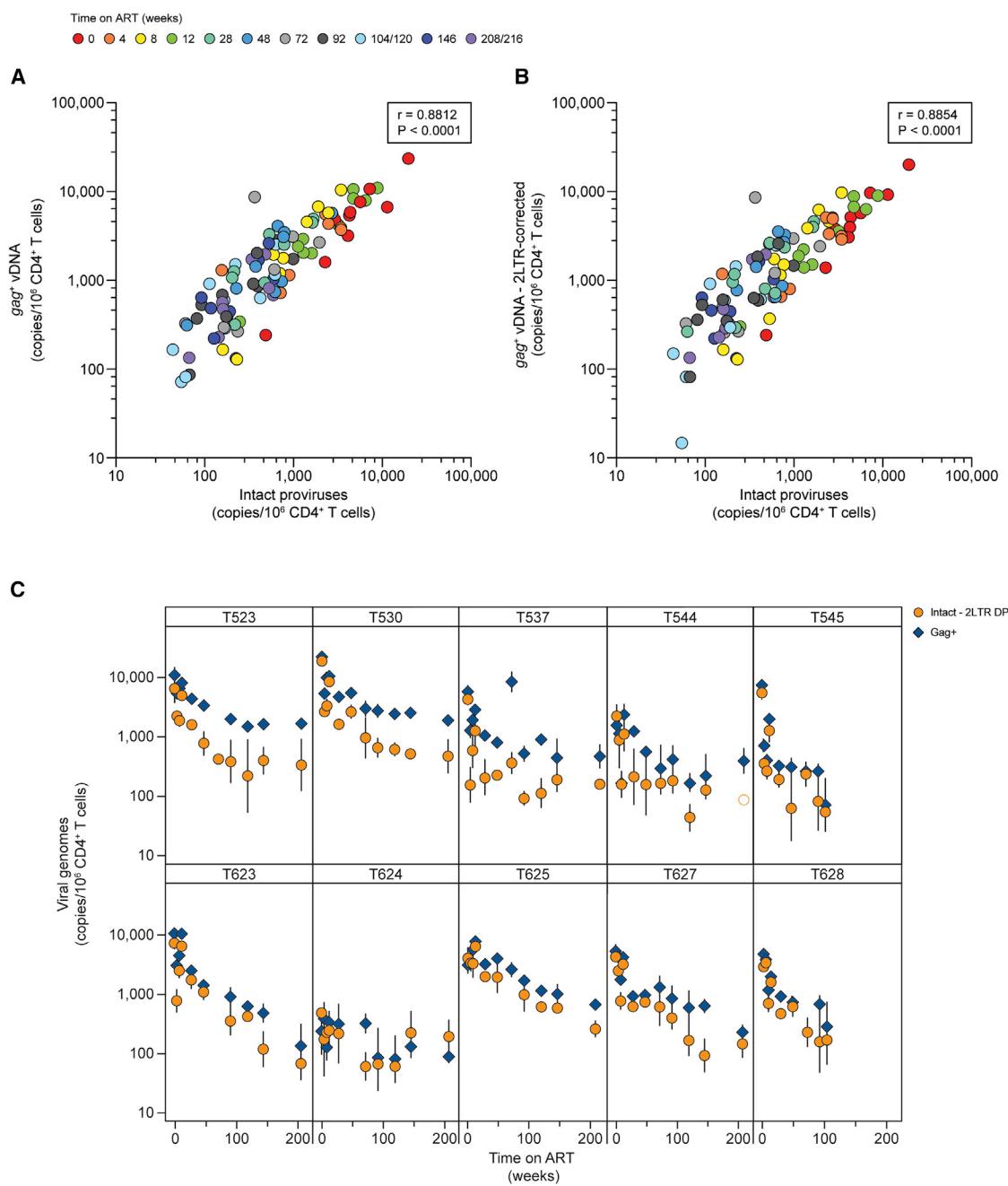


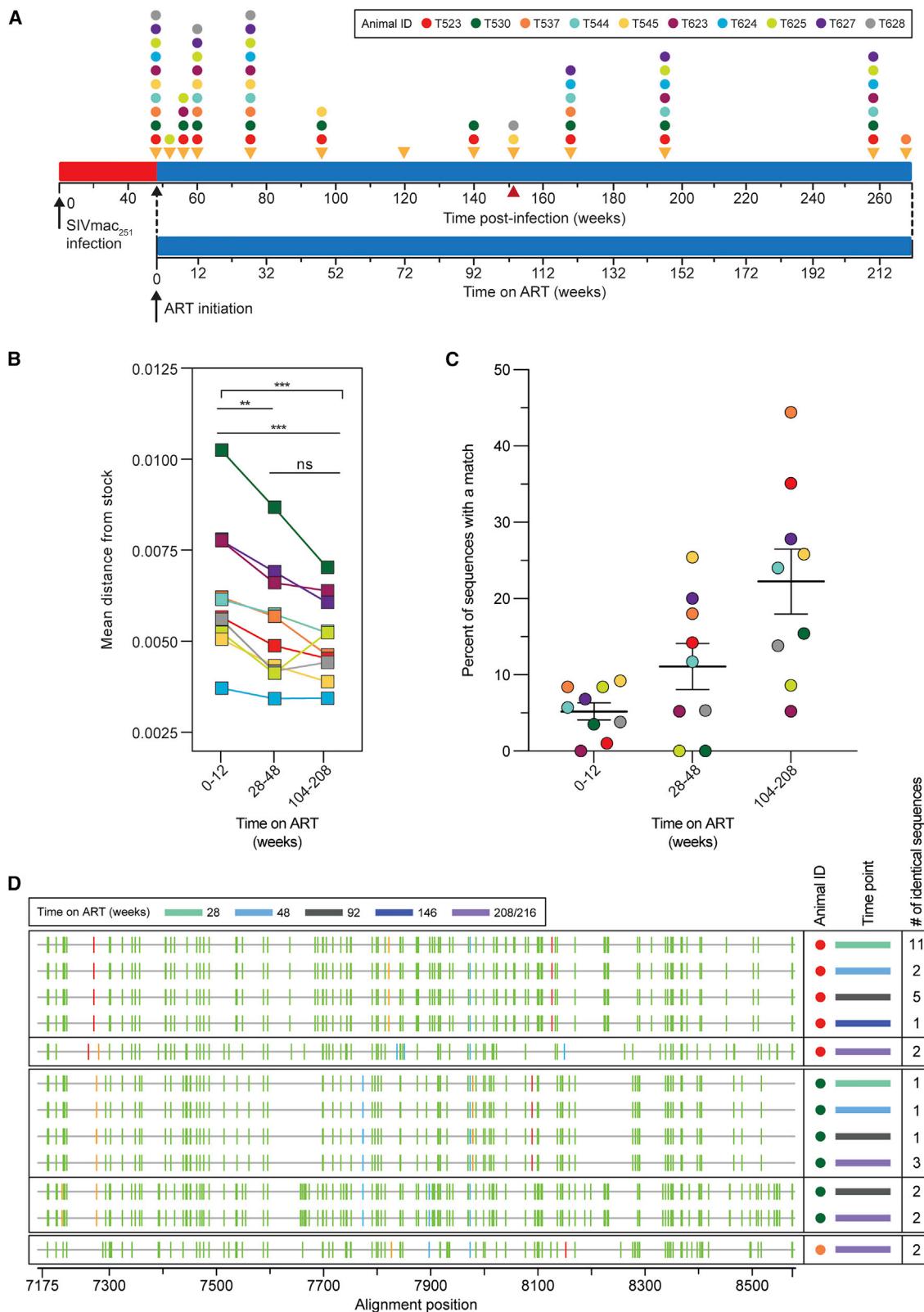
Figure 4. Comparison of *gag* ddPCR and IPDA quantification of SIV DNA

(A) Correlation between *gag** SIV DNA copies and intact proviruses quantified using the IPDA. Individual datapoints represent the number of copies detected using either an amplicon in *gag* (y-axis)⁶⁵ or the IPDA (x-axis).²⁰ Each point is the geometric mean of 3 replicates. Intact proviruses are corrected for shearing (DSI) and *env**2LTR circles. The correlation between the two variables was calculated using Pearson's coefficient. Data from one animal, T624, were excluded from this analysis due to failure of the *gag* amplicon resulting from mutations or deletion.

(B) Correlation between *gag** SIV DNA copies and intact proviruses, with *gag* values corrected using the same *env**2LTR correction factor applied to the IPDA data, calculated using Pearson's coefficient.

(C) Comparison of the decay of SIV *gag* copies and intact proviruses for the animals in cohort 18-02. IPDA data are corrected for *env**2LTR circles and DNA shearing. Vertical lines represent the standard deviations.

See also Figure S4.



(legend on next page)

sequences collected from 3 animals (Figure 5D) and captured identical sequences. Proviruses with extensive hypermutation are unable to replicate²⁰ and therefore must arise through proliferation or through infection of multiple cells by the same viral variant with APOBEC3G/F-mediated hypermutation subsequently occurring in exactly the same way in each cell. Analysis of the likelihood of G→A mutation at each G in *env* (see STAR Methods) showed that the likelihood of the latter scenario is extremely low ($p < 10^{-7}$). Thus, infected cells with identical hypermutated *env* sequences arise through a proliferative process that also likely explains at least some of the identical intact *env* sequences observed at later time points.

env sequence analysis and evidence of antibody pressure

We also compared *env* sequences from plasma virus at ART initiation to non-defective proviral *env* sequences from circulating CD4⁺ T cells at various time points on ART (Figure 5A). A representative maximum likelihood phylogenetic tree for animal T530 is shown in Figure 6. This tree is rooted on the consensus of the stock sequences⁵⁶ and illustrates the enormous sequence divergence from the stock and intra-individual diversification arising during untreated infection. Similar patterns were observed for all other animals (Figures S5–S12). Although all animals were infected with the same stock, the divergence occurred differently in each animal as shown by a composite phylogenetic tree (Figure S13A). The exception is T624 (Figure S13B), which spontaneously controlled viral replication to low levels before ART (Figure S1). However, for the other 9 animals, in all cases mutations accumulated in the gp120 variable regions. The V1/V2 and V4 regions of the SIV *env* gene are under strong selective pressure from the antibody response,^{74–88} and key amino acid positions consistently accumulate non-synonymous mutations in MHC-disparate animals, coincident with the development of autologous neutralizing antibodies (aNabs). We found many previously reported mutations,^{78,86} including at N202, W345, D415, R420, P421, and R424. In 9 animals, we found the R424Q mutation, an early indicator of escape from humoral immunity.⁸⁶ Other mutations alter potential N-linked glycosylation sites (PNGs), including D415N, which creates a new PNG, and non-synonymous changes at residue N202, which removes a PNG in V1.^{81,82,86} These mutations were observed in more than one, and sometimes as many as 9 of 10 animals. In 2 animals (T530 and T623), we also observed characteristic deletions in V4 (Figures 6 and S9) that mediate escape from aNabs.^{78,86}

Plasma sequences at ART initiation show high divergence from the stock relative to proviral sequences in CD4⁺ T cells

The time 0 plasma sequences represent virus actively replicating at ART initiation. In addition to the ~1,700 sequences collected from CD4⁺ T cell DNA, we collected over 300 plasma sequences from 9 animals at week 0. Compared with proviral *env* sequences in circulating CD4⁺ T cells at the same time, plasma sequences showed a high mean divergence from the stock in all 9 animals analyzed (Figure 7A). Plasma sequences were consistently mutated at the positions associated with antibody pressure (Figures 6 and S5–S12). No wild-type sequences that clustered closer to the root were recovered from plasma from any animal at week 0 (Figures 6, 7B, and S5–S12). Importantly, the time 0 plasma sequences were found predominantly on distal branches (Figures 6, 7B, and S5–S12). This can be illustrated by comparing the number of non-synonymous mutations away from the stock per sequence for the time 0 plasma and cellular sequences (Figure 7B). Some circulating CD4⁺ T cells present at ART initiation contained SIV variants that had accumulated antibody escape mutations similar to those in plasma, while other CD4⁺ T cells present at the same time harbored variants that had fewer of these mutations and that were not seen in plasma at the time of ART initiation. These are likely ancestral sequences from earlier pre-ART time points preserved in the latent reservoir. This finding was consistent for all animals and not related to the depth of sampling (Figure 7B).

Sequences with large numbers of antibody escape mutations become less abundant at later time points

In contrast to the normal pattern of increasing divergence over time in untreated infection,^{40,72,86,89,90} sequences with large numbers of antibody escape mutations became less abundant at the later time points on ART (Figure 7C). This is evident in the phylogenetic trees (Figures 6 and S5–S12). Relatively unmutated sequences close to the root were frequently isolated at later time points. At weeks 120–216 on ART, sequences with few or no mutations relative to the stock became more apparent. We analyzed the kinetics of the decay process by examining changes in the mean number of non-synonymous mutations per sequence (Figure 7D). Modeling revealed a biphasic decay pattern. The initial phase of decay has a half-life of 11.4 weeks (95% CI: 4.6–28.1 weeks). The second phase is so slow as to be statistically indistinguishable from “no decay.” The net result of these qualitative changes is summarized in Figure 7C, which shows that at almost all positions in the *env* ectodomain mutations are less frequent after 3 years of ART than at ART initiation.

Figure 5. Analysis of *env* sequences from CD4⁺ T cell DNA during ART

- (A) Timeline for sequence analysis. Colored dots indicate sampling times for each animal.
- (B) Average pairwise distance from the stock⁵⁶ of *env* sequences collected at early (weeks 0–12), intermediate (weeks 28–48) or late (weeks 104–216) time points after ART initiation. The number of sequences analyzed for each animal is listed in Table S7. Statistical significance was determined using one-way ANOVA (GraphPad Prism).
- (C) Percent of non-defective *env* sequences that exactly match another independent sequence from the same sample at early, intermediate, or late time points on ART. Sequences with hypermutation (as determined using Los Alamos National Lab’s Hypermut Program⁹⁷), or deletions were excluded from this analysis. Animal T624 was excluded due to a high frequency of identical sequences at all time points, reflecting early control of replication (Figures 1B and S13).
- (D) Highlighter plots of identical hypermutated *env* sequences from 3 animals. The animal (colored dot, see key in A), time point (colored bar), and the number of times an identical sequence was recovered from a given sample are indicated. Vertical bars indicate a nucleotide difference from the consensus of the stock.⁵⁶ The color of the bar represents the mutant nucleotide (red, thymine; blue, cytosine; orange, guanine; green, adenine). Alignment numbering corresponds to nucleotide positions based on the SIVmac₂₃₉ reference (accession no: M33262.1).

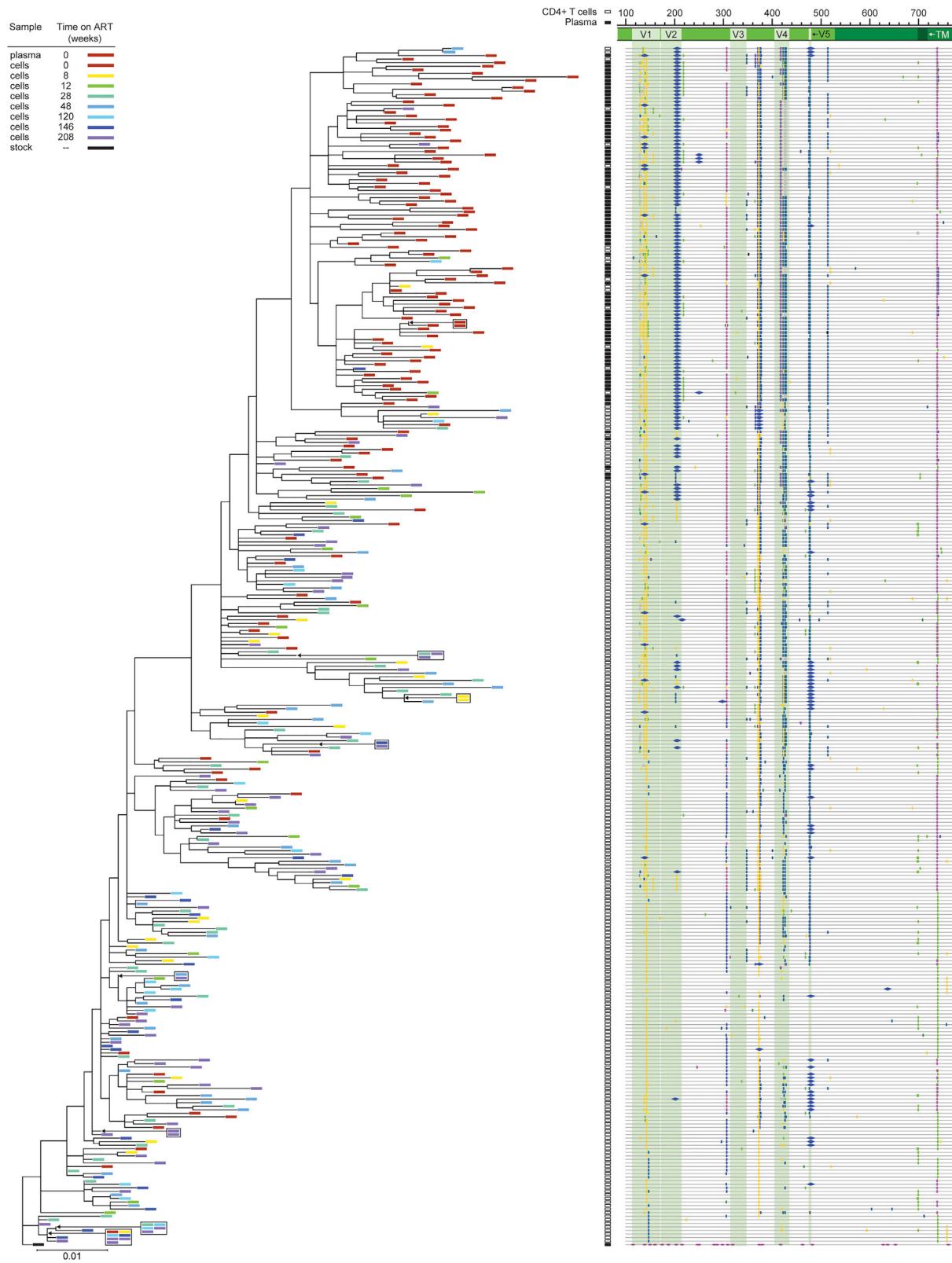


Figure 6. Sequencing analysis of *env* diversity in plasma virus and proviral DNA from CD4⁺ T cells for T530

Phylogenetic tree and highlighter plot of *env* sequences collected from either plasma or CD4⁺ T cell DNA from animal T530. Highlighter plot was generated using Los Alamos National Lab's Highlighter program. The tree is rooted on a consensus sequence of the stock.⁵⁶ Sequence analysis includes 2,054 nt.

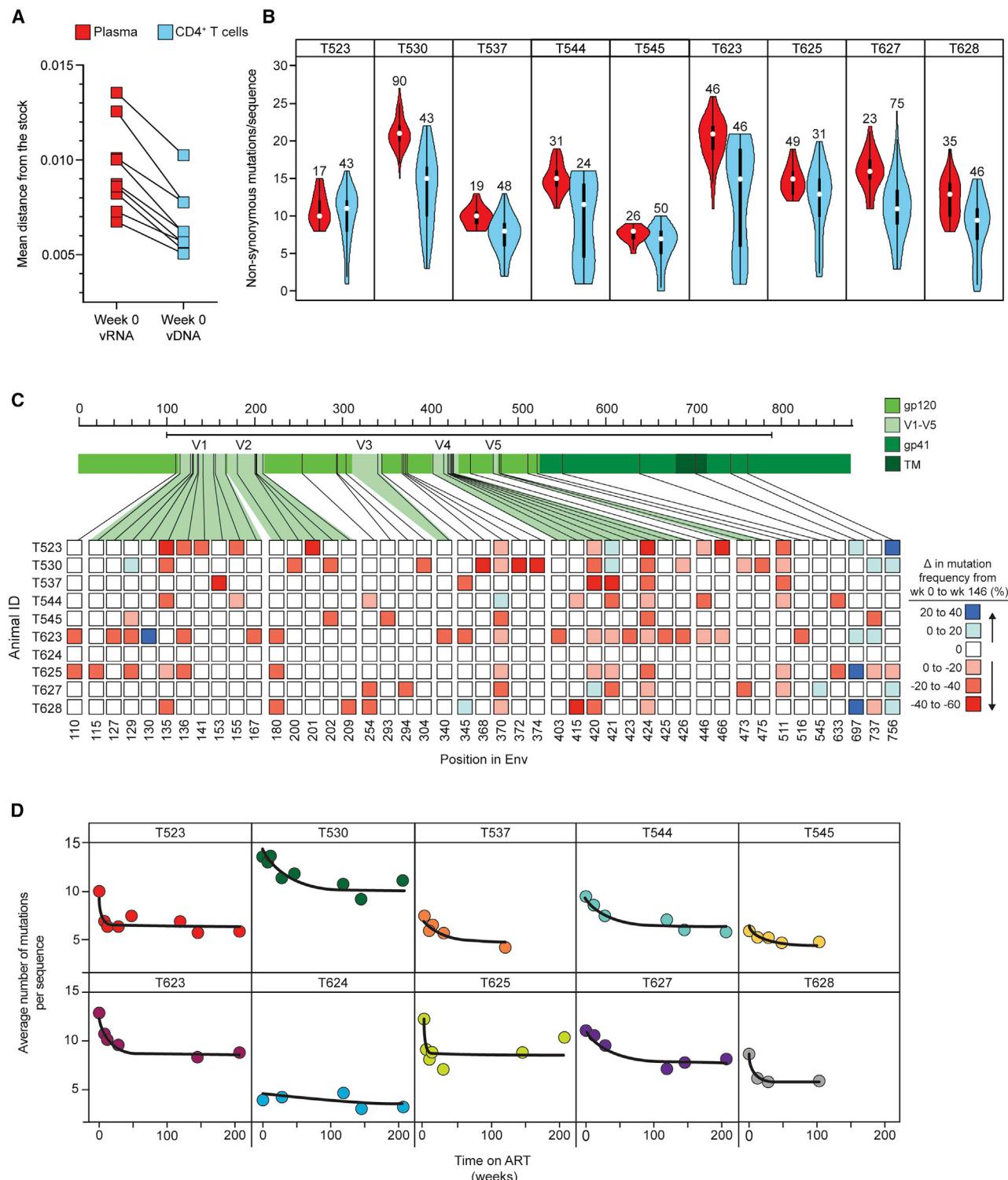


Figure 7. Differences in mutation frequency between time 0 plasma virus and cellular sequences at time 0 and week 146

(A) Comparison of average pairwise distance from the stock⁵⁶ between sequences collected from plasma (red) or peripheral blood CD4⁺ T cell DNA (blue) at the time of ART initiation ($t = 0$). The number of sequences analyzed, and standard deviations are given in Table S7.

(B) Violin plots comparing the number of non-synonymous amino acid mutations per *env* sequence in plasma (red) vs. CD4⁺ T cell DNA (blue) at the time of ART initiation. Numbers above the violin plots indicate the number of sequences included in the analysis.

(C) Heatmap depicting changes in mutation frequency at different positions in *env* between sequences collected from week 0 and week 146 for each animal. Week 146 was chosen as a late time point because slight increases in mutational frequency were observed between week 146 and 208, reflecting distortions

(legend continued on next page)

This finding is consistent with the decay process described above. The infected cells that decay after ART initiation carry proviruses with multiple immune escape mutations. As these cells decay, less mutated ancestral sequences persisting in the latent reservoir become more apparent. Strikingly, many of these variants closest to the root of the trees are from later on ART time points (Figures 6 and S5–S12). These results provide strong evidence that ART arrests viral evolution, but more importantly, they show that viruses are deposited in the latent reservoir throughout untreated infection and not just at ART initiation.

DISCUSSION

To inform HIV-1 cure strategies, we characterized quantitative and qualitative changes in SIV-infected CD4⁺ T cells following ART initiation during chronic infection and over the next 4 years. Our findings provide critical insights into the dynamics of infected cells over a time frame in which many intervention studies are conducted.

Overall, the decay kinetics of cells harboring intact SIV proviruses are similar to those reported for HIV-1.²⁵ In both cases, there is complex, multiphasic decay, with the final phase representing the stable persistence of a latent reservoir in CD4⁺ T cells. For both viruses, there are different decay rates for intact and defective proviruses, likely reflecting selective pressures operating on cells capable of expressing viral genes. In both cases, cells with 2LTR circles decay with kinetics similar to those of intact proviruses. Together, these findings support the use of the non-human primate (NHP) model in HIV-1 cure research.

Our study did uncover important differences in decay kinetics that must be considered in cure studies. Notably, although the frequency of intact genomes at the start of ART was similar for SIV and HIV-1 (geometric mean frequency = 3,642 vs. 2,255 per 10⁶ CD4⁺ T cells,²⁵ the first and second phase half-lives for intact SIV genomes are significantly shorter than those reported for HIV-1 (3.3 days and 8.1 months for SIV vs. 12.9 days and 19 months for HIV-1²⁵). Differences in the second phase decay kinetics are particularly important given that most cure studies in NHP models involve interventions carried out during second phase decay. The relatively quasi-stable third phase is not reached until ~2.3 years of ART. Studies of cure interventions during this second phase must include robust untreated controls to distinguish natural vs. intervention-induced decay.

By following treated macaques for 4 years, we observed in the same animals both the early and late phase dynamics shaping the latent reservoir. Most of the plasma virus is produced by rapidly decaying cells that are not present in the circulation (Figures S1D and S1E).⁵⁹ However, most infected CD4⁺ T cells in the circulation at ART initiation are also labile and decay during the first 2.3 years without entering the stable latent reservoir. The stable reservoir becomes apparent as rapidly decaying cells are eliminated (Figures 5, 6, and 7). The proliferation of infected cells

generates large clones, which after several years comprise a substantial fraction of the reservoir (Figures 5C and 5D). This proliferation is particularly evident in HIV-1-infected individuals on long-term ART^{66–68} and may be responsible for the inflection in the decay of intact proviruses seen after 7 years of ART.²⁴

We also developed an assay for hypermutated SIV proviruses. Most hypermutated SIV genomes have premature stop codons in most genes^{20,91} and are unlikely to give rise to high-level viral gene expression. We demonstrate slower decay of cells with hypermutated proviruses relative to intact proviruses, consistent with selective processes acting against cells that express viral genes.^{21,23–25,92}

We also compared a single amplicon in *gag* to the IPDA. The *gag* assay provides a reasonable measure of intact proviruses at the beginning of ART. However, it becomes progressively less accurate with increasing time on ART as cells with intact proviruses decay and cells with defective proviruses, many of which are detected by the *gag* assay, persist.

Lastly, we used *env* SGS to determine whether the quantitative changes in infected cell populations were accompanied by qualitative changes. We show that infected cell populations change in composition during ART. The populations that decay initially are recently infected cells carrying proviruses with immune escape mutations. As these populations decay, ancestral variants with fewer mutations that are archived in the latent reservoir and that are not present in the plasma at the time of ART initiation become more prominent. Our results show that sequences arising throughout the period of untreated infection can be deposited in the latent reservoir. The reservoir includes both early sequences with few escape mutations as well as sequences that have accumulated more escape mutations and are deposited shortly before the time of ART initiation. All of these variants can persist in the reservoir without immune selection as long as viruses remain latent. What changes is the relative proportion of these early and late variants. The change is due to the decay processes defined here. At the time of ART initiation, the population of productively infected cells, as sampled by sequencing of plasma virus, consists exclusively of sequences with many escape mutations (red bars associated with black squares in Figures 6 and S5–S12). Some of these variants persist and are deposited in the reservoir after several years of ART (blue and purple bars near the top of Figures 6 and S5–S12). Variants with fewer mutations are also preserved in the reservoir and become more prominent with increasing time on ART, as most of the variants that were actively replicating at the time of ART initiation decay. In the long-term, variants from both early and late in untreated infection persist in the reservoir.

This finding is important in the context of recent studies suggesting that the reservoir is predominantly comprised cells infected at or around the time of ART initiation.^{51,52,55,93} The observed decrease in APD from the stock, the observation that sequences with more mutations in the *env* variable regions

introduced by the proliferation of particular infected cell clones. Red or blue squares indicate increases or decreases, respectively, in the frequency of sequences with a mutation at the indicated position. Analysis covers sequenced positions where there were mutations away from the stock consensus in at least 20% of sequences at that time point.

(D) Decay of non-synonymous mutations in *env* sequences during ART. Data were fitted to a biphasic model (black line) as described in the STAR Methods section. The number of sequences included in this analysis, mean APD, and standard deviation for each animal at the indicated time points are included in Table S7.

become less apparent at later time points, and the emergence of archival variants during ART are more consistent with a model in which sequences enter the reservoir continually throughout untreated infection.⁵³ Latency explains how the cells harboring archival sequences persist during untreated infection in the context of high levels of immune activation, CD4⁺ T cell death and surveillance from the host immune system. Some of these factors are altered by ART, skewing the reservoir toward variants replicating around the time of treatment initiation. However, it is incorrect to conclude that the reservoir is composed exclusively of such variants.

There are limitations to our study. First, we analyzed only peripheral blood CD4⁺ T cells. Although we hypothesize that the increase in intact proviruses observed at week 12 is due to the redistribution of CD4⁺ T cells from tissues to the blood, we have not directly tested this due to limited tissue sample availability from this longitudinal cohort.

Furthermore, the discrepancy between the initial decay rates of plasma viremia and intact proviruses in the blood suggests that the cells producing plasma virus are found at another anatomical location. It will be important to test this hypothesis in future studies using contemporaneous blood and tissue samples. Another important consideration is that the animals in cohort 18–02 were treated during chronic infection. Many NHP studies initiate ART much earlier and may show different decay dynamics.

Together, these findings reveal that the population of SIV-infected CD4⁺ T cells is dynamic and changes quantitatively and qualitatively for several years following ART initiation. These data provide a framework for evaluating and interpreting intervention trials utilizing the SIV/NHP model.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Animals
- METHOD DETAILS
 - SIV viral load quantification
 - CD4⁺ T cell isolation & DNA extraction
 - Droplet digital PCR (ddPCR)
 - RPP30
 - IPDA
 - env-2LTRc
 - Hypermutated Proviral DNA Assay (HPDA)
 - HPDA – sensitivity and specificity
 - HPDA – sequence validation
 - gag ddPCR
 - gag mutation analysis
 - Envelope sequencing & analysis
 - Phylogenetic analysis
 - Isolation of RNA from plasma for sequencing

- Modeling of decay processes
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.chom.2023.01.016>.

ACKNOWLEDGMENTS

This work was supported by the NIH Martin Delaney Collaboratory grant UM1AI164556 to D.H.B. and by Howard Hughes Medical Institute to R.F.S. Animal studies were supported by AI124377, AI128751, AI149670, AI164556, and AI169615 for D.H.B. F.R.S is supported by the Office of the NIH Director and National Institute of Dental and Craniofacial Research (DP5OD031834). Portions of this work were performed under the auspices of the US Department of Energy through Los Alamos National Laboratory, which is operated by Triad National Security, LLC, for the National Nuclear Security Administration of the US Department of Energy (contract 89233218CNA000001). Support was also provided by NIH grants R01-AI028433, R01-OD011095, P01-AI131365 (to A.S.P.), R01-AI15270301, and UM1-AI164561 (to R.M.R.).

AUTHOR CONTRIBUTIONS

E.J.F., R.F.S., J.D.S., and D.H.B. designed the study. E.J.F., R.F.S., J.D.S., D.H.B., J.D.V., and F.W., F.R.S., C.Z., N.S., C.M.P., T.L., R.M.R., and A.S.P. participated in discussion and interpretation of the results. E.J.F. and F.W. performed the experiments. A.M.B. provided samples that were used for assay development and optimization. D.H.B., P.L., J.D.V., and R.G. provided samples, information, and drugs related to the animal work. C.Z., N.S., C.M.P., T.L., R.M.R., A.S.P., performed the mathematical modeling. R.W. and D.H.O. performed MHC genotyping and analysis. E.J.F. wrote the manuscript. R.F.S. and J.D.S. revised the manuscript. revised the manuscript. E.J.F., F.R.S., C.Z., N.S., T.L., R.M.R., A.S.P., and R.F.S. performed the data analysis

DECLARATION OF INTERESTS

Aspects of HIV-1 IPDA are subject of a patent application PCT/US16/28822 filed by Johns Hopkins University. R.F.S. is an inventor on this application. Accelevir Diagnostics holds an exclusive license for this patent application. R.F.S. holds no equity interest in Accelevir Diagnostics.

Received: September 19, 2022

Revised: December 2, 2022

Accepted: January 24, 2023

Published: February 20, 2023

REFERENCES

1. Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., and Ho, D.D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* 387, 188–191.
2. Chun, T.W., Stuyver, L., Mizell, S.B., Ehler, L.A., Mican, J.A., Baseler, M., Lloyd, A.L., Nowak, M.A., and Fauci, A.S. (1997). Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc. Natl. Acad. Sci. USA* 94, 13193–13197.
3. Finzi, D., Blankson, J., Siliciano, J.D., Margolick, J.B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., et al. (1999). 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.
4. Wong, J.K., Hezareh, M., Günthard, H.F., Havlir, D.V., Ignacio, C.C., Spina, C.A., and Richman, D.D. (1997). Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* 278, 1291–1295.

5. Crooks, A.M., Bateson, R., Cope, A.B., Dahl, N.P., Griggs, M.K., Kuruc, J.D., Gay, C.L., Eron, J.J., Margolis, D.M., Bosch, R.J., and Archin, N.M. (2015). Precise quantitation of the latent HIV-1 reservoir: implications for eradication strategies. *J. Infect. Dis.* 212, 1361–1365.
6. Siliciano, J.D., Kajdas, J., Finzi, D., Quinn, T.C., Chadwick, K., Margolick, J.B., Kovacs, C., Gange, S.J., and Siliciano, R.F. (2003). 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.
7. Davey, R.T., Bhat, N., Yoder, C., Chun, T.W., Metcalf, J.A., Dewar, R., Natarajan, V., Lempicki, R.A., Adelsberger, J.W., Miller, K.D., et al. (1999). 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. USA* 96, 15109–15114.
8. Rothenberger, M.K., Keele, B.F., Wietgrefe, S.W., Fletcher, C.V., Beilman, G.J., Chipman, J.G., Khoruts, A., Estes, J.D., Anderson, J., Callisto, S.P., et al. (2015). Large number of rebounding/founder HIV variants emerge from multifocal infection in lymphatic tissues after treatment interruption. *Proc. Natl. Acad. Sci. USA* 112, E1126–E1134.
9. Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., and Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373, 123–126.
10. Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., and Hahn, B.H. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373, 117–122.
11. Bruner, K.M., Murray, A.J., Pollack, R.A., Soliman, M.G., Laskey, S.B., Capoferi, A.A., Lai, J., Strain, M.C., Lada, S.M., Hoh, R., et al. (2016). Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nat. Med.* 22, 1043–1049.
12. Cohn, L.B., Silva, I.T., Oliveira, T.Y., Rosales, R.A., Parrish, E.H., Learn, G.H., Hahn, B.H., Czartoski, J.L., McElrath, M.J., Lehmann, C., et al. (2015). HIV-1 integration landscape during latent and active infection. *Cell* 160, 420–432.
13. Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S., and Malim, M.H. (2003). DNA deamination mediates innate immunity to retroviral infection. *Cell* 113, 803–809.
14. Hiener, B., Horsburgh, B.A., Eden, J.S., Barton, K., Schluß, T.E., Lee, E., von Stockenstrom, S., Odevall, L., Milush, J.M., Liegler, T., et al. (2017). Identification of genetically intact HIV-1 proviruses in specific CD4+ T cells from effectively treated participants. *Cell Rep.* 21, 813–822.
15. Ho, Y.C., Shan, L., Hosmane, N.N., Wang, J., Laskey, S.B., Rosenbloom, D.I.S., Lai, J., Blankson, J.N., Siliciano, J.D., and Siliciano, R.F. (2013). Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* 155, 540–551.
16. Imamichi, H., Dewar, R.L., Adelsberger, J.W., Rehm, C.A., O'Doherty, U., Paxinos, E.E., Fauci, A.S., and Lane, H.C. (2016). Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. *Proc. Natl. Acad. Sci. USA* 113, 8783–8788.
17. Lee, G.Q., Orlova-Fink, N., Einkauf, K., Chowdhury, F.Z., Sun, X., Harrington, S., Kuo, H.H., Hua, S., Chen, H.R., Ouyang, Z., et al. (2017). Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. *J. Clin. Invest.* 127, 2689–2696.
18. Sanchez, G., Xu, X., Chermann, J.C., and Hirsch, I. (1997). Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. *J. Virol.* 71, 2233–2240.
19. Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral vif protein. *Nature* 418, 646–650.
20. Bender, A.M., Simonetti, F.R., Kumar, M.R., Fray, E.J., Bruner, K.M., Timmons, A.E., Tai, K.Y., Jenike, K.M., Antar, A.A.R., Liu, P.T., et al. (2019). The landscape of persistent viral genomes in ART-treated SIV, SHIV, and HIV-2 infections. *Cell Host Microbe* 26, 73–85.e4.
21. Bruner, K.M., Wang, Z., Simonetti, F.R., Bender, A.M., Kwon, K.J., SenGupta, S., Fray, E.J., Beg, S.A., Antar, A.A.R., Jenike, K.M., et al. (2019). A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature* 566, 120–125.
22. Falcinelli, S.D., Kilpatrick, K.W., Read, J., Murtagh, R., Allard, B., Ghofrani, S., Kirchherr, J., James, K.S., Stuelke, E., Baker, C., et al. (2021). Longitudinal dynamics of intact HIV proviral DNA and outgrowth virus frequencies in a cohort of individuals receiving antiretroviral therapy. *J. Infect. Dis.* 224, 92–100.
23. Gandhi, R.T., Cytor, J.C., Bosch, R.J., Mar, H., Laird, G.M., Martin, A., Collier, A.C., Riddler, S.A., Macatangay, B.J., Rinaldo, C.R., et al. (2021). Selective decay of intact HIV-1 proviral DNA on antiretroviral therapy. *J. Infect. Dis.* 223, 225–233.
24. Peluso, M.J., Bacchetti, P., Ritter, K.D., Beg, S., Lai, J., Martin, J.N., Hunt, P.W., Henrich, T.J., Siliciano, J.D., Siliciano, R.F., et al. (2020). Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy. *JCI Insight* 5, e123997.
25. White, J.A., Simonetti, F.R., Beg, S., McMyn, N.F., Dai, W., Bachmann, N., Lai, J., Ford, W.C., Bunch, C., Jones, J.L., et al. (2022). Complex decay dynamics of HIV virions, intact and defective proviruses, and 2LTR circles following initiation of antiretroviral therapy. *Proc. Natl. Acad. Sci. USA* 119, e2120326119.
26. Bukrinsky, M.I., Stanwick, T.L., Dempsey, M.P., and Stevenson, M. (1991). Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection. *Science* 254, 423–427.
27. Zack, J.A., Arrigo, S.J., Weitsman, S.R., Go, A.S., Haislip, A., and Chen, I.S. (1990). HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. *Cell* 61, 213–222.
28. Pierson, T.C., Zhou, Y., Kieffer, T.L., Ruff, C.T., Buck, C., and Siliciano, R.F. (2002). Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection. *J. Virol.* 76, 8518–8531.
29. Zhou, Y., Zhang, H., Siliciano, J.D., and Siliciano, R.F. (2005). Kinetics of human immunodeficiency virus type 1 decay following entry into resting CD4+ T cells. *J. Virol.* 79, 2199–2210.
30. Murray, J.M., Kelleher, A.D., and Cooper, D.A. (2011). Timing of the components of the HIV life cycle in productively infected CD4+ T cells in a population of HIV-infected individuals. *J. Virol.* 85, 10798–10805.
31. Butler, S.L., Johnson, E.P., and Bushman, F.D. (2002). Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. *J. Virol.* 76, 3739–3747.
32. Pierson, T.C., Kieffer, T.L., Ruff, C.T., Buck, C., Gange, S.J., and Siliciano, R.F. (2002). Intrinsic stability of episomal circles formed during human immunodeficiency virus type 1 replication. *J. Virol.* 76, 4138–4144.
33. Sharkey, M.E., Teo, I., Greenough, T., Sharova, N., Luzuriaga, K., Sullivan, J.L., Bucy, R.P., Kostrikis, L.G., Haase, A., Veryard, C., et al. (2000). Persistence of episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nat. Med.* 6, 76–81.
34. Sharkey, M., Triques, K., Kuritzkes, D.R., and Stevenson, M. (2005). In vivo evidence for instability of episomal human immunodeficiency virus type 1 cDNA. *J. Virol.* 79, 5203–5210.
35. Del Prete, G.Q., and Lifson, J.D. (2018). Nonhuman primate models for studies of AIDS virus persistence during suppressive combination antiretroviral therapy. *Curr. Top. Microbiol. Immunol.* 417, 69–109.
36. Dinozo, J.B., Rabi, S.A., Blankson, J.N., Gama, L., Mankowski, J.L., Siliciano, R.F., Zink, M.C., and Clements, J.E. (2009). A simian immunodeficiency virus-infected macaque model to study viral reservoirs that persist during highly active antiretroviral therapy. *J. Virol.* 83, 9247–9257.

37. Kumar, N., Chahroudi, A., and Silvestri, G. (2016). Animal models to achieve an HIV cure. *Curr. Opin. HIV AIDS* 11, 432–441.
38. Zhang, Z., Schuler, T., Zupancic, M., Wietgrefe, S., Staskus, K.A., Reimann, K.A., Reinhart, T.A., Rogan, M., Cavert, W., Miller, C.J., et al. (1999). Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells. *Science* 286, 1353–1357.
39. Brandin, E., Thorstensson, R., Bonhoeffer, S., and Albert, J. (2006). Rapid viral decay in simian immunodeficiency virus-infected macaques receiving quadruple antiretroviral therapy. *J. Virol.* 80, 9861–9864.
40. Keele, B.F., Li, H., Learn, G.H., Hraber, P., Giorgi, E.E., Grayson, T., Sun, C., Chen, Y., Yeh, W.W., Letvin, N.L., et al. (2009). Low-dose rectal inoculation of rhesus macaques by SIVsmE660 or SIVmac251 recapitulates human mucosal infection by HIV-1. *J. Exp. Med.* 206, 1117–1134.
41. Li, Q., Duan, L., Estes, J.D., Ma, Z.M., Rourke, T., Wang, Y., Reilly, C., Carlis, J., Miller, C.J., and Haase, A.T. (2005). Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* 434, 1148–1152.
42. Ma, Z.M., Abel, K., Rourke, T., Wang, Y., and Miller, C.J. (2004). A period of transient viremia and occult infection precedes persistent viremia and antiviral immune responses during multiple low-dose intravaginal simian immunodeficiency virus inoculations. *J. Virol.* 78, 14048–14052.
43. Malouli, D., Gilbride, R.M., Wu, H.L., Hwang, J.M., Maier, N., Hughes, C.M., Newhouse, D., Morrow, D., Ventura, A.B., Law, L., et al. (2022). Cytomegalovirus-vaccine-induced unconventional T cell priming and control of SIV replication is conserved between primate species. *Cell Host Microbe* 30, 1207–1218.e7.
44. Nowak, M.A., Lloyd, A.L., Vasquez, G.M., Wiltzout, T.A., Wahl, L.M., Bischofberger, N., Williams, J., Kinter, A., Fauci, A.S., Hirsch, V.M., et al. (1997). Viral dynamics of primary viremia and antiretroviral therapy in simian immunodeficiency virus infection. *J. Virol.* 71, 7518–7525.
45. Okoye, A.A., Hansen, S.G., Vaidya, M., Fukazawa, Y., Park, H., Duell, D.M., Lum, R., Hughes, C.M., Ventura, A.B., Ainslie, E., et al. (2018). Early antiretroviral therapy limits SIV reservoir establishment to delay or prevent post-treatment viral rebound. *Nat. Med.* 24, 1430–1440.
46. Shen, A., Zink, M.C., Mankowski, J.L., Chadwick, K., Margolick, J.B., Carruth, L.M., Li, M., Clements, J.E., and Siliciano, R.F. (2003). Resting CD4+ T lymphocytes but not thymocytes provide a latent viral reservoir in a simian immunodeficiency virus-macaca nemestrina model of human immunodeficiency virus type 1-infected patients on highly active antiretroviral therapy. *J. Virol.* 77, 4938–4949.
47. Swanson, A.E., Del Prete, G.Q., Deleage, C., Elser, S.E., Lackner, A.A., and Hoxie, J.A. (2018). The SIV envelope glycoprotein, viral tropism, and pathogenesis: novel insights from nonhuman primate models of AIDS. *Curr. HIV Res.* 16, 29–40.
48. Veazey, R.S., DeMaria, M., Chalifoux, L.V., Shvetz, D.E., Pauley, D.R., Knight, H.L., Rosenzweig, M., Johnson, R.P., Desrosiers, R.C., and Lackner, A.A. (1998). Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. *Science* 280, 427–431.
49. Whitney, J.B., Hill, A.L., Sanisetty, S., Penalosa-MacMaster, P., Liu, J., Shetty, M., Parenteau, L., Cabral, C., Shields, J., Blackmore, S., et al. (2014). Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature* 512, 74–77.
50. Long, S., Fennessey, C.M., Newman, L., Reid, C., O'Brien, S.P., Li, Y., Del Prete, G.Q., Lifson, J.D., Gorelick, R.J., and Keele, B.F. (2019). Evaluating the intactness of persistent viral genomes in simian immunodeficiency virus-infected rhesus macaques after initiating antiretroviral therapy within one year of infection. *J. Virol.* 94, e01308-19.
51. Abrahams, M.R., Joseph, S.B., Garrett, N., Tyers, L., Moeser, M., Archin, N., Council, O.D., Matten, D., Zhou, S., Doolabh, D., et al. (2019). The replication-competent HIV-1 latent reservoir is primarily established near the time of therapy initiation. *Sci. Transl. Med.* 11, eaaw5589.
52. Brodin, J., Zanini, F., Thebo, L., Lanz, C., Bratt, G., Neher, R.A., and Albert, J. (2016). Establishment and stability of the latent HIV-1 DNA reservoir. *eLife* 5, e18889.
53. Brooks, K., Jones, B.R., Dilernia, D.A., Wilkins, D.J., Claiborne, D.T., McNally, S., Gilmour, J., Kilembe, W., Joy, J.B., Allen, S.A., et al. (2020). HIV-1 variants are archived throughout infection and persist in the reservoir. *PLoS Pathog.* 16, e1008378.
54. Joos, B., Fischer, M., Kuster, H., Pillai, S.K., Wong, J.K., Böni, J., Hirscher, B., Weber, R., Trkola, A., and Günthard, H.F.; Swiss HIV Cohort Study (2008). HIV rebounds from latently infected cells, rather than from continuing low-level replication. *Proc. Natl. Acad. Sci. USA* 105, 16725–16730.
55. Pankau, M.D., Reeves, D.B., Harkins, E., Ronen, K., Jaoko, W., Mandaliya, K., Graham, S.M., McClelland, R.S., Matsen, F.A., Schiffer, J.T., et al. (2020). Dynamics of HIV DNA reservoir seeding in a cohort of superinfected Kenyan women. *PLoS Pathog.* 16, e1008286.
56. Del Prete, G.Q., Scarlotta, M., Newman, L., Reid, C., Parodi, L.M., Roser, J.D., Oswald, K., Marx, P.A., Miller, C.J., Desrosiers, R.C., et al. (2013). Comparative characterization of transfection- and infection-derived simian immunodeficiency virus challenge stocks for in vivo nonhuman primate studies. *J. Virol.* 87, 4584–4595.
57. Del Prete, G.Q., Smedley, J., Macallister, R., Jones, G.S., Li, B., Hattersley, J., Zheng, J., Piatak, M., Keele, B.F., Hesselgesser, J., et al. (2016). Short communication: comparative evaluation of coformulated injectable combination antiretroviral therapy regimens in simian immunodeficiency virus-infected rhesus macaques. *AIDS Res. Hum. Retroviruses* 32, 163–168.
58. Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M., and Ho, D.D. (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 271, 1582–1586.
59. He, X., Aid, M., Ventura, J.D., Borducchi, E., Lifton, M., Liu, J., and Barouch, D.H. (2022). Rapid loss of CD4 T cells by pyroptosis during acute SIV infection in rhesus macaques. *J. Virol.* 96, e0080822.
60. Pollicchio, B.B., Cardozo, E.F., Sette, P., Xu, C., Haret-Richter, G., Dunsmore, T., Apetrei, C., Pandrea, I., and Ribeiro, R.M. (2018). Dynamics of simian immunodeficiency virus two-long-terminal-repeat circles in the presence and absence of CD8+ cells. *J. Virol.* 92, e02100-17.
61. Bucy, R.P., Hockett, R.D., Derdeyn, C.A., Saag, M.S., Squires, K., Sillers, M., Mitsuyasu, R.T., and Kilby, J.M. (1999). Initial increase in blood CD4(+) lymphocytes after HIV antiretroviral therapy reflects redistribution from lymphoid tissues. *J. Clin. Invest.* 103, 1391–1398.
62. Pakker, N.G., Notermans, D.W., de Boer, R.J., Roos, M.T., de Wolf, F., Hill, A., Leonard, J.M., Danner, S.A., Miedema, F., and Schellekens, P.T. (1998). Biphasic kinetics of peripheral blood T cells after triple combination therapy in HIV-1 infection: a composite of redistribution and proliferation. *Nat. Med.* 4, 208–214.
63. Lifson, J.D., Rossio, J.L., Piatak, M., Parks, T., Li, L., Kiser, R., Coalter, V., Fisher, B., Flynn, B.M., Czajak, S., et al. (2001). Role of CD8(+) lymphocytes in control of simian immunodeficiency virus infection and resistance to rechallenge after transient early antiretroviral treatment. *J. Virol.* 75, 10187–10199.
64. Mattapallil, J.J., Douek, D.C., Hill, B., Nishimura, Y., Martin, M., and Roederer, M. (2005). Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature* 434, 1093–1097.
65. Cline, A.N., Bess, J.W., Piatak, M., and Lifson, J.D. (2005). Highly sensitive SIV plasma viral load assay: practical considerations, realistic performance expectations, and application to reverse engineering of vaccines for AIDS. *J. Med. Primatol.* 34, 303–312.
66. Bui, J.K., Sobolewski, M.D., Keele, B.F., Spindler, J., Musick, A., Wiegand, A., Luke, B.T., Shao, W., Hughes, S.H., Coffin, J.M., et al. (2017). Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS Pathog.* 13, e1006283.
67. Hosmane, N.N., Kwon, K.J., Bruner, K.M., Capoferri, A.A., Beg, S., Rosenbloom, D.I., Keele, B.F., Ho, Y.C., Siliciano, J.D., and Siliciano, R.F. (2017). Proliferation of latently infected CD4+ T cells carrying

- replication-competent HIV-1: potential role in latent reservoir dynamics. *J. Exp. Med.* 214, 959–972.
68. Lorenzi, J.C., Cohen, Y.Z., Cohn, L.B., Kreider, E.F., Barton, J.P., Learn, G.H., Oliveira, T., Lavine, C.L., Horwitz, J.A., Settler, A., et al. (2016). Paired quantitative and qualitative assessment of the replication-competent HIV-1 reservoir and comparison with integrated proviral DNA. *Proc. Natl. Acad. Sci. USA* 113, E7908–E7916.
 69. Maldarelli, F., Wu, X., Su, L., Simonetti, F.R., Shao, W., Hill, S., Spindler, J., Ferris, A.L., Mellors, J.W., Kearney, M.F., et al. (2014). HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* 345, 179–183.
 70. Wagner, T.A., McLaughlin, S., Garg, K., Cheung, C.Y.K., Larsen, B.B., Styrchak, S., Huang, H.C., Edlefsen, P.T., Mullins, J.I., and Frenkel, L.M. (2014). HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science* 345, 570–573.
 71. Ferris, A.L., Wells, D.W., Guo, S., Del Prete, G.Q., Swanstrom, A.E., Coffin, J.M., Wu, X., Lifson, J.D., and Hughes, S.H. (2019). Clonal expansion of SIV-infected cells in macaques on antiretroviral therapy is similar to that of HIV-infected cells in humans. *PLoS Pathog.* 15, e1007869.
 72. Laskey, S.B., Pohlmeier, C.W., Bruner, K.M., and Siliciano, R.F. (2016). Evaluating clonal expansion of HIV-infected cells: optimization of PCR strategies to predict clonality. *PLoS Pathog.* 12, e1005689.
 73. Patro, S.C., Brandt, L.D., Bale, M.J., Halvas, E.K., Joseph, K.W., Shao, W., Wu, X., Guo, S., Murrell, B., Wiegand, A., et al. (2019). Combined HIV-1 sequence and integration site analysis informs viral dynamics and allows reconstruction of replicating viral ancestors. *Proc. Natl. Acad. Sci. USA* 116, 25891–25899.
 74. Buckley, K.A., Li, P.L., Khimani, A.H., Hofmann-Lehmann, R., Liska, V., Anderson, D.C., McClure, H.M., and Ruprecht, R.M. (2003). Convergent evolution of SIV env after independent inoculation of rhesus macaques with infectious proviral DNA. *Virology* 312, 470–480.
 75. Burns, D.P., and Desrosiers, R.C. (1991). Selection of genetic variants of simian immunodeficiency virus in persistently infected rhesus monkeys. *J. Virol.* 65, 1843–1854.
 76. Campbell, B.J., and Hirsch, V.M. (1994). Extensive envelope heterogeneity of simian immunodeficiency virus in tissues from infected macaques. *J. Virol.* 68, 3129–3137.
 77. Choi, W.S., Collignon, C., Thiria, C., Burns, D.P., Stott, E.J., Kent, K.A., and Desrosiers, R.C. (1994). Effects of natural sequence variation on recognition by monoclonal antibodies neutralize simian immunodeficiency virus infectivity. *J. Virol.* 68, 5395–5402.
 78. Ita, S., Hill, A.K., Lam, E.C., Dufort, F.J., Yang, X., Newman, R., Leviyang, S., Fofana, I.B., and Johnson, W.E. (2018). High-resolution sequencing of viral populations during early simian immunodeficiency virus infection reveals evolutionary strategies for rapid escape from emerging env-specific antibody responses. *J. Virol.* 92, e01574–e17.
 79. Kinsey, N.E., Anderson, M.G., Unangst, T.J., Joag, S.V., Narayan, O., Zink, M.C., and Clements, J.E. (1996). Antigenic variation of SIV: mutations in V4 alter the neutralization profile. *Virology* 221, 14–21.
 80. Lamers, S.L., Nolan, D.J., Rife, B.D., Fogel, G.B., McGrath, M.S., Burdo, T.H., Autissier, P., Williams, K.C., Goodnow, M.M., and Salemi, M. (2015). Tracking the emergence of host-specific simian immunodeficiency virus env and nef populations reveals nef early adaptation and convergent evolution in brain of naturally progressing rhesus macaques. *J. Virol.* 89, 8484–8496.
 81. Overbaugh, J., Rudensey, L.M., Papenhausen, M.D., Benveniste, R.E., and Morton, W.R. (1991). Variation in simian immunodeficiency virus env is confined to V1 and V4 during progression to simian AIDS. *J. Virol.* 65, 7025–7031.
 82. Overbaugh, J., and Rudensey, L.M. (1992). Alterations in potential sites for glycosylation predominate during evolution of the simian immunodeficiency virus envelope gene in macaques. *J. Virol.* 66, 5937–5948.
 83. Sato, S., Yuste, E., Lauer, W.A., Chang, E.H., Morgan, J.S., Bixby, J.G., Lifson, J.D., Desrosiers, R.C., and Johnson, W.E. (2008). Potent antibody-mediated neutralization and evolution of antigenic escape variants of simian immunodeficiency virus strain SIVmac239 *in vivo*. *J. Virol.* 82, 9739–9752.
 84. Torres, J.V., Anderson, D.E., Malley, A., Banapour, B., Axthelm, M.K., Benjamini, E., and Gardner, M.B. (1993). SIV envelope glycoprotein epitopes recognized by antibodies from infected or vaccinated rhesus macaques. *J. Med. Primatol.* 22, 129–137.
 85. Wu, F., Ourmanov, I., Kuwata, T., Goeken, R., Brown, C.R., Buckler-White, A., Iyengar, R., Plishka, R., Aoki, S.T., and Hirsch, V.M. (2012). Sequential evolution and escape from neutralization of simian immunodeficiency virus SIVsmE660 clones in rhesus macaques. *J. Virol.* 86, 8835–8847.
 86. Yeh, W.W., Rahman, I., Hraber, P., Coffey, R.T., Nevidomskyte, D., Giri, A., Asmal, M., Mijikovic, S., Daniels, M., Whitney, J.B., et al. (2010). Autologous neutralizing antibodies to the transmitted/founder viruses emerge late after simian immunodeficiency virus SIVmac251 infection of rhesus monkeys. *J. Virol.* 84, 6018–6032.
 87. Yeh, W.W., Brassard, L.M., Miller, C.A., Basavapathruni, A., Zhang, J., Rao, S.S., Nabel, G.J., Mascola, J.R., Letvin, N.L., and Seaman, M.S. (2012). Envelope variable region 4 is the first target of neutralizing antibodies in early simian immunodeficiency virus mac251 infection of rhesus monkeys. *J. Virol.* 86, 7052–7059.
 88. Yuste, E., Johnson, W., Pavlakis, G.N., and Desrosiers, R.C. (2005). Virion envelope content, infectivity, and neutralization sensitivity of simian immunodeficiency virus. *J. Virol.* 79, 12455–12463.
 89. Keele, B.F., Giorgi, E.E., Salazar-Gonzalez, J.F., Decker, J.M., Pham, K.T., Salazar, M.G., Sun, C., Grayson, T., Wang, S., Li, H., et al. (2008). Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc. Natl. Acad. Sci. USA* 105, 7552–7557.
 90. Shankarappa, R., Margolick, J.B., Gange, S.J., Rodrigo, A.G., Upchurch, D., Farzadegan, H., Gupta, P., Rinaldo, C.R., Learn, G.H., He, X., et al. (1999). Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J. Virol.* 73, 10489–10502.
 91. Liu, P.T., Keele, B.F., Abbink, P., Mercado, N.B., Liu, J., Bondzie, E.A., Chandrashekhar, A., Borducchi, E.N., Hesselgesser, J., Mish, M., et al. (2020). Origin of rebound virus in chronically SIV-infected rhesus monkeys following treatment discontinuation. *Nat. Commun.* 11, 5412.
 92. Einkauf, K.B., Osborn, M.R., Gao, C., Sun, W., Sun, X., Lian, X., Parsons, E.M., Gladkov, G.T., Seiger, K.W., Blackmer, J.E., et al. (2022). Parallel analysis of transcription, integration, and sequence of single HIV-1 proviruses. *Cell* 185, 266–282.e15.
 93. Jones, B.R., Kinloch, N.N., Horacek, J., Ganase, B., Harris, M., Harrigan, P.R., Jones, R.B., Brockman, M.A., Joy, J.B., Poon, A.F.Y., and Brumme, Z.L. (2018). Phylogenetic approach to recover integration dates of latent HIV sequences within-host. *Proc. Natl. Acad. Sci. USA* 115, E8958–E8967.
 94. Wiseman, R.W., Karl, J.A., Bimber, B.N., O'Leary, C.E., Lank, S.M., Tuscher, J.J., Detmer, A.M., Bouffard, P., Levenkova, N., Turcotte, C.L., et al. (2009). Major histocompatibility complex genotyping with massively parallel Pyrosequencing. *Nat. Med.* 15, 1322–1326.
 95. Barouch, D.H., Ghneim, K., Bosche, W.J., Li, Y., Berkemeier, B., Hull, M., Bhattacharyya, S., Cameron, M., Liu, J., Smith, K., et al. (2016). Rapid inflamasome activation following mucosal SIV infection of rhesus monkeys. *Cell* 165, 656–667.
 96. Stone, M., Keele, B.F., Ma, Z.M., Bailes, E., Dutra, J., Hahn, B.H., Shaw, G.M., and Miller, C.J. (2010). A limited number of simian immunodeficiency virus (SIV) env variants are transmitted to rhesus macaques vaginally inoculated with SIVmac251. *J. Virol.* 84, 7083–7095.

97. Rose, P.P., and Korber, B.T. (2000). Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. *Bioinformatics* 16, 400–401.
98. Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027.
99. Tosiano, M.A., Jacobs, J.L., Shutt, K.A., Cykota, J.C., and Mellors, J.W. (2019). A simpler and more sensitive single-copy HIV-1 RNA assay for quantification of persistent HIV-1 viremia in individuals on suppressive antiretroviral therapy. *J. Clin. Microbiol.* 57, e01714-18.
100. Burnham, K.P., and Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, (Second Edition) (Springer).
101. Lavielle, M. (2014). *Mixed Effects Models for the Population Approach: Models, Tasks, Methods and Tools* (Chapman and Hall/CRC).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
SIVmac ₂₅₁ swarm ⁵⁶	Dr. Dan Barouch, BIDMC	N/A
Biological samples		
SIVmac ₂₅₁ -infected Rhesus macaque PBMCs	Dr. Dan Barouch, BIDMC	N/A
Chemicals, peptides, and recombinant proteins		
Platinum Taq DNA Polymerase High Fidelity	ThermoFisher Scientific	Cat # 11304011
SuperScript III Reverse Transcriptase	ThermoFisher Scientific	Cat # 18-080-044
dNTP Mix (10 mM each)	ThermoFisher Scientific	Cat # 18-427-088
UltraPure 1M Tris-HCl Buffer (pH 8.0)	ThermoFisher Scientific	Cat # 15-568-025
RNaseOUT Recombinant Ribonuclease Inhibitor	ThermoFisher Scientific	Cat # 10777019
Critical commercial assays		
QIAamp DNA Mini Kit	Qiagen	Cat # 51304
CD4+ T Cell Isolation Kit, Non-Human Primate	Miltenyi Biotec	Cat # 130-092-144
2X ddPCR Supermix for Probes (no dUTPs)	Bio-Rad	Cat # 186-3024
Deposited data		
SIV – plasma RNA sequences	This paper	Genbank: OQ168641-OQ168979
SIV – non-defective proviral DNA sequences	This paper	Genbank: OQ168980-OQ170751
SIV – hypermutated (defective) proviral sequences	This paper	Genbank: OQ170752-OQ170785
Experimental models: Organisms/strains		
Rhesus macaque (<i>Macaca mulatta</i>) infected with SIVmac ₂₅₁	Indian origin	Animals T523, T530, T537, T544, T545, T623, T624, T625, T627 & T628
Oligonucleotides		
Oligonucleotides, standard desalting (Sanger sequencing & ddPCR) – see STAR Methods and Table S8	IDT	N/A
FAM/VIC labeled probes with MGB quencher, HPLC purified (see Table S8)	Applied Biosystems	N/A
Unlabeled competition probes w/MGB quencher, HPLC purified (see Table S8)	Applied Biosystems	N/A
FAM/HEX labeled probes w/MGB-NFQ quencher, HPLC purified (see Table S8)	IDT	N/A
Recombinant DNA		
Synthetic double-stranded DNA controls (gBlocks) – see STAR Methods	IDT	N/A
Software and algorithms		
QuantaSoft Analysis-Pro	Bio-Rad	http://www.bio-rad.com/en-us/ product/qx200-droplet-digital- pcr-system?ID=MPOQQE4VY
Geneious Prime	Dotmatics	https://www.geneious.com/prime/
MEGA v7.0 & v11	MEGA	https://www.megasoftware.net/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Clustal W	EMBL-EBI	http://www.ebi.ac.uk/Tools/msa/clustalw2/
GraphPad Prism	Dotmatics	https://www.graphpad.com/
Monolix 2020R1	Lixoft	https://lixoft.com/
Los Alamos National Labs – Highlighter	Keele et al. ⁸⁹	https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html
Los Alamos National Labs – Hypermut 2.0	Rose et al. ⁹⁷	https://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Robert F. Siliciano (rsiliciano@jhmi.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Sequences generated in this study have been deposited in Genbank and are publicly available as of the date of publication. Accession numbers are listed in the [key resources table](#).
- This study does not report original code and all computational resources used are publicly available and listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

Animal work was approved by the Institutional Care & Use Committees of Bioqual and the NIH and determined to be in accordance with the guidelines outlined by the Animal Welfare Act and Regulation (USDA) and the Guide for the Care & Use of Laboratory Animals, 8th Edition (NIH).

Cohort 15-09: The details of this study have been described elsewhere.^{20,91} Near-full-length sequencing data collected from SIV-infected Rhesus macaques (n=7, Benderet al.²⁰; or n=16, Liu et al.⁹¹) were used for assay development and verification of assay targets, as detailed in the “HPDA sequence validation” subsection below.

Cohort 18-02: 10 outbred Indian-origin rhesus macaques (*Macaca mulatta*) were housed at Bioqual Inc. in Rockville, MD. Animals were infected via repetitive intrarectal challenge with the SIVmac₂₅₁ swarm⁵⁶ until infection was confirmed by detection of SIV RNA in the plasma via qPCR. After ~48 weeks, the animals were placed on a daily regimen of tenofovir disoproxil fumarate, emtricitabine, and dolutegravir (TDF/FTC/DTG, Gilead Sciences, Inc.⁵⁷) The animals were MHC class I genotyped as previously described.⁹⁴ MHC genotypes, ages and sexes of the animals in this cohort are detailed below.

Animal ID	Sex	Age at time of infection (years)	MHC-A haplotype 1	MHC-A haplotype 2	MHC-B haplotype 1	MHC-B haplotype 2
T523	F	7.6	A*002	A*701	B*017	B*017
T530	F	6.6	A*023	A*002	B*001	B*024
T537	F	5.8	A*011	A*012	B*012	B*012
T544	M	3.6	A*004	A*019	B*024	B*015
T545	M	4.8	A*004	A*002	B*015	B*055
T623	M	4.9	A*004	A*018	B*015	B*076
T624	M	5.1	A*028	A*002	B*001	B*055
T625	M	4.8	A*002	A*002	B*015	B*043
T627	M	3.3	A*004	A*025	B*048	B*069
T628	M	4.4	A*001	A*004	B*012	

METHOD DETAILS**SIV viral load quantification**

Plasma SIV RNA was quantified using a qRT-PCR assay targeting *gag* as described.⁹⁵

CD4⁺ T cell isolation & DNA extraction

PBMCs were collected from whole blood, viably frozen in DMSO, and stored in liquid nitrogen until sample processing. Cells were rapidly thawed at 37°C in RPMI + 20% heat-inactivated FBS and subjected to negative depletion using the Miltenyi NHP CD4 Isolation Kit according to the manufacturer's protocol. After negative depletion, up to 5×10^6 cells were pelleted and resuspended in 200 µl PBS, then lysed for DNA extraction using the QiAamp DNA Mini Kit according to the manufacturer's directions.

Droplet digital PCR (ddPCR)

Specific primer and probe sequences for each assay are described in [Table S8](#).^{20,65} Details specific to individual assays are described below under sub-headings. Conditions common to all ddPCR experiments are described here:

Samples were assayed in triplicate unless indicated otherwise. Reactions were conducted in a total volume of 22 µl with 10 µl 2X Bio-Rad ddPCR Supermix (no dUTPs). Primers were included at a final concentration of 600 nM, and probes at a final concentration of 200 nM. Droplets were made using the Bio-Rad QX200 manual droplet generator and thermal-cycled using a Bio-Rad QX200 Thermalcycler and the following conditions: 95°C for 10 minutes, 50 cycles of (30 seconds at 95°C, 2 minutes at 56°C), and 10 minutes at 98°C, with a ramp rate of 2°C/s. Cycling conditions were modified from the original protocol described by Bender et al.²⁰ to include more cycles, a lower extension temperature and a longer extension time. Droplets were read by the Bio-Rad QX200 Reader and analyzed with Quantasoft Studio Software. Wells with <10,000 droplets were not analyzed.

RPP30

Host genomes in each sample were quantified using a duplex ddPCR assay with 2 amplicons specific for the Rhesus macaque Ribonuclease P/MRP Subunit 30 (*RPP30*) gene and spaced at the same distance as the SIV IPDA amplicons.^{20,21} The *RPP30* primer/probe pairs are listed in [Table S8](#). Primers were included at a final concentration of 500 nM and probes at a final concentration of 250 nM. *RPP30* data were used to correct for DNA shearing and to determine the number of host cells per sample.

IPDA

The IPDA was performed as described²⁰ using the primer/probe pairs indicated in [Table S8](#).

env-2LTRc

The *env*-2LTRc duplex assay was performed as described^{20,60} using the primer/probe pairs indicated in [Table S8](#).

Hypermutated Proviral DNA Assay (HPDA)

The HPDA was performed using the primer/probe pairs described in [Table S8](#). 6 replicates were performed per sample to account for the overall lower frequency of hypermutated proviruses. Details of assay design, optimization and validation are given below.

HPDA – sensitivity and specificity

The HPDA is an assay designed to quantify hypermutated proviruses, based on the same principles as the IPDA ([Figure S3A](#)).¹ To test whether the HPDA correctly distinguishes between templates that are intact at both positions and those which are hypermutated at both positions, we designed synthetic DNA templates to match the *pol* and *env* sequence encompassing the region of the HPDA amplicons ([Figure S3B](#)). We designed these templates to reflect the intact probe sequence (no G→A mutations) or hypermutated probe sequence (two G→A mutations) at one or both amplicons, and tested them using the HPDA, or, as a control, the IPDA ([Figure S3C](#)). Both assays performed as predicted for proviruses matching hypermutated or intact sequences at one or both amplicons ([Figure S3C](#)). These data are consistent with control experiments performed when validating the SIV IPDA¹ and provide evidence that switching the labeled and unlabeled probes does not affect the assay's ability to differentiate between intact and hypermutated templates. Using these controls, we determined that the HPDA is sensitive and specific for hypermutated proviruses and performs well using the same thermal-cycling conditions as the IPDA.

HPDA – sequence validation

Although the HPDA performed well when used to measure templates which exactly match one of two target sequences, *in vivo*, SIV proviruses may contain polymorphisms which could affect probe binding. Because the assay we describe here essentially relies on two bases within each amplicon to discriminate intact and hypermutated proviruses, it is important to determine whether additional sequence variation in the probe site could affect our ability to detect hypermutated genomes in biological samples from infected macaques. We therefore analyzed the same 266 sequences used to design the IPDA¹ for sequence conservation within the probe binding site ([Figures S3D](#) and [S3E](#)). By this analysis, approximately 30% of hypermutated proviruses would be missed by our assay due to deletions encompassing both probe sites. However, ~90% of proviruses with sequence information for at least one of two probe sites (i.e. proviruses without deletions at both amplicons) would be detected by the HPDA. In addition to these 266 sequences,

we extended our analysis to a dataset of 259 near full-length (nFL) hypermutated sequences collected from 16 SIVmac₂₅₁-infected macaques from that same cohort (Figures S3F and S3G) (15–09).³ This dataset was generated using PCR product size as an exclusion-criteria to specifically collect nFL sequences, therefore proviruses with deletions are not included. The majority (81% for *pol* and 66% for *env*) of hypermutated proviruses included in this analysis directly matched the HPDA probe sequences, emphasizing that the HPDA probe positions with G→A mutations are indeed good indicators of APOBEC3G/F-mediated hypermutation status. Interestingly, proviruses with additional G→A mutations (three or four) in the *env* probe binding site were relatively common in this dataset (~11%, compared to ~3% in the first dataset). Of the nFL hypermutated sequences with 3 or 4 G→A mutations in the region of the HPDA *env* probe, 93% had 2 G→A mutations at the *pol* probe binding site and would therefore be detected by the HPDA in Q1. Although approximately 16% of the hypermutated sequences from the second dataset had SNPs that could prevent the hypermutated *env* probe from binding, we found that 85% of these sequences were hypermutated at both positions in the *pol* probe binding site. Based on this analysis, we concluded that the most accurate way to estimate the frequency of hypermutated proviruses would be to sum the positive events detected in quadrants 1, 2 and 4 by the HPDA to give an estimate of “total hypermutated proviruses,” rather than attempt to categorize positive droplets based on the region of hypermutation (i.e. 5'HM (Q1), 3'HM (Q4) or 2XHM (Q2)). An important implication of this analysis is that single-positive events do not necessarily represent proviruses that are deleted at the other amplicon.

gag ddPCR

We designed synthetic double-stranded DNA templates (corresponding to nt: 1185–1620 in the SIVmac₂₅₁ reference genome – accession no: M19499.1) to adapt the previously described SIV *gag* qPCR assay⁶⁵ to the Bio-Rad ddPCR platform (Figures S4A and S4C; Table S8). Primers were included at a final concentration of 500 nM and the probe at a final concentration of 250 nM.

gag mutation analysis

Sequences from SIV-infected macaques^{20,91} were analyzed for G→A sequence polymorphisms in the *gag* probe binding site (Figure S4C). Sequences with 1-to-5 mutations were used to design synthetic DNA templates to determine the effect of mismatches on assay performance, including the amplitude of positive droplets (Figures S4A and S4C). 250 copies of each template were loaded in a 22 μl reaction in triplicate and scored for fluorescence in the FAM channel.

Envelope sequencing & analysis

DNA used for ddPCR was also used to collect individual *env* sequences by limiting dilution nested PCR. Each sample was diluted for an outer PCR using the primers SOUTF (5' – GGC TAA GGC TAA TAC ATC TTC TGC ATC – 3') and NouTR (5' – TTT AAG CAA GCA AGC GTG GAG – 3')⁸⁰ at a final concentration of 0.2 μM with Platinum Taq and the following cycling conditions: 5 minutes at 95°C, 35 cycles of (94°C for 1 minute, 58°C for 1 minute, 72°C for 3 minutes) then 10 minutes at 72°C.⁸⁰ PCR products were diluted 1 in 2 with 10 mM Tris-HCl and 2 μl were used as the template for a nested inner PCR with the primers EnvF2 (5' – GGA ACA ACT CAG TGC CTA CCA GAT AAT GGT G – 3') and EnvR2 (5' – GTA GGT CAG TTC AGT CCT GAG GAC TTC TCG – 3')⁹⁶ at a final concentration of 0.20 μM and the following PCR conditions: 2 minutes at 94°C, 41 cycles of (94°C for 30 seconds, 55°C for 30 seconds, 72°C for 2 minutes) then 3 minutes at 72°C.⁹⁶ 3 μl of each inner PCR product was added to 17 μl loading dye, run on an agarose gel and visualized using UV illumination. Gels in which <30% of wells were positive were considered to be at limiting dilution and to have resulted from a single template. Positive wells were submitted for Sanger Sequencing (Genewiz). Chromatograms were assembled using the “de novo assembly” function in Geneious Prime. Chromatograms with one or more overlapping peaks – indicative of Taq polymerase-induced error or multiple proviral templates in a single well – were removed from analysis.

Phylogenetic analysis

Sequences were defined as “non-defective” based on the absence of premature stop codons, deletions, or evidence of APOBEC3G/F-mediated G-to-A hypermutation, assessed using Los Alamos National Lab’s Hypermut 2.0 program.⁹⁷ Only non-defective sequences were used for phylogenetic tree construction (Figures 6 and S5–S13) and other analyses including: percent of sequences with a match (Figure 5), average pairwise-distance calculations (Table S7; Figures 5B and 7A), and frequency of non-synonymous mutations (Figures 7B–7D). The evolutionary history was inferred by using the Maximum Likelihood method implemented in MEGA11⁹⁸ using the optimal model selected by the program. Trees were rooted using a consensus of the stock sequences.⁵⁶

Average pairwise distances were calculated between the consensus sequence of the virus stock⁵⁶ and all non-defective envelope nucleotide sequences retrieved from plasma and gDNA. Calculations were implemented in MEGA v.7 with a p-distance substitution model. Analyses were run for 1000 replicates.

To determine the likelihood of identical hypermutated sequences arising independently in different infected cells, we calculated the frequency of G→A mutation at each G in the sequenced region of *env* in 226 hypermutated sequences from 3 animals (Figure 5D) and used per site mutations probabilities in simulations to determine how often identical hypermutated sequences would arise assuming each site behaved independently.

Isolation of RNA from plasma for sequencing

To allow for consistent amplification of plasma virus, plasma aliquots were thawed and ~10,000 copies of SIV RNA were used. Viral RNA was extracted as previously described.⁹⁹ Viral RNA pellets were resuspended in 20μl Buffer TDR (5 mM Tris-HCl, 1 mM DDT,

2.5 mM RNaseOUT) and immediately used for reverse transcription with SuperScript III. A mixture of 20 µl SIV RNA, 2.5 µl of 2 µM primer NOUTR (5' – TTT AAG CAA GCA AGC GTG GAG – 3')⁸⁰ and 2.5 µl 10 mM dNTPs was heated to 65°C for 10 minutes then placed on ice for 1 minute. A reaction mixture of 10 µl 5X First Strand Buffer, 13.5 µl water, 0.5 µl 0.1 M DTT, 0.5 µl SuperScript III and 0.5 µl 40 U/µl RNaseOUT was then added to reaction volume of 50 µl. The mixture was heated at 50°C for 55 minutes then inactivated at 85°C for 10 minutes. The resulting cDNA was serially diluted with 10 mM Tris-HCl to reach limiting dilution as input for single-genome sequencing using the approach described above.

Modeling of decay processes

We used a non-linear mixed effect approach to fit multiple models for the decay of SIV DNA species in macaques infected with SIV on ART. Given the long time scale of the study, the following general model was used:

$$y = \begin{cases} y_0(Ae^{-b_1 t} + (1 - A)e^{-b_2 t}) & \text{if } t \leq T_s \\ y_0(Ae^{-b_1 T_s} + (1 - A)e^{-b_2 T_s})e^{-b_3 t} & \text{if } t > T_s \end{cases} \quad (\text{Equation 1})$$

where y is the variable of interest, y_0 is its baseline value, A is the fraction of y that decays in the first phase with decay rate b_1 , and $(1 - A)$ is the fraction of y which decays in the second phase with decay rate b_2 . After this potential biphasic decay lasting until time T_s , there could be a third phase with decay rate b_3 , which would be smaller than the two initial phases.

With this general formula, multiple models can be tested. If we fix $b_3=0$, then the third phase of decay starting at time T_s is flat and remains at the value of y attained at time T_s . On the other hand, if we fix T_s at a large value (say $T_s=10,000$) then we never reach T_s and have the usual biphasic decay model. In this latter case, with $A = 0$, we can test for a single decay model, estimating only b_2 . Note that in the case of a single decay, b_1 is not estimated. Using this approach, we statistically compare the quality of the fits (using the corrected Bayesian Information Criteria, BIC) among these models.¹⁰⁰ In addition, we also tested a hybrid model, (particularly for the hypermutated DNA), where some macaques (specifically T544, T545, T624, T627 and T628) had a biphasic decay, and another group of macaques (T523, T530, T537, T623 and T625) had a single phase of decay. This was prompted by visual inspection of the decay patterns across the animals.

To fit the model of the decay processes – equations (1) above – to the data, we used a nonlinear mixed-effects model. This approach assumes that a model parameter η_i is drawn from a distribution with a fixed part θ , which is the median value of the parameter in the population, and a random term φ_i , which is assumed to be normally distributed with zero mean and standard deviation σ . We assumed that parameters b_1 , b_2 , b_3 follow a lognormal distribution (to guarantee that they are positive), and A is logit normal (between 0 and 1). We fit instances of the model to the data, with different parameters fixed as described above, and estimate the parameters using the software Monolix 2020R1 (Lixoft).¹⁰¹ We calculated 95% CI for the parameters based on the Fisher Information matrix using the Monolix function confintmlx.

For intact proviruses, *gag*⁺ DNA and 2LTR circles, the best fit models were triphasic, with the last phase after T_s , essentially flat, $b_3=0$. For the *env*⁺ and *env*⁻ 2LTR species, the biphasic and triphasic models had similar statistical support, with biphasic slightly better, because T_s was very high. We suspect that this is the case due to multiple censored values (i.e. no positive events were detected) for these species. We note that we couldn't fit the 2LTR species for T624, because it showed almost no decay with many censored values. Finally, for hypermutated SIV proviruses, the hybrid model gave the best fits.

We used the same general modeling equations (1) to quantify the decay of non-synonymous mutations over time, which was well described by a biphasic decline.

QUANTIFICATION AND STATISTICAL ANALYSIS

Details regarding the statistical tests used and the represented data can be found in the figure legends, or, for the mathematical modeling in the “Modeling of decay processes” section. Statistical tests were considered significant when P values were less than 0.05. Statistical analyses were performed via GraphPad Prism (Prism Software) or, for the modeling of decay processes, Monolix2020R1 (Lixoft).

Supplemental information

**Antiretroviral therapy reveals triphasic
decay of intact SIV genomes
and persistence of ancestral variants**

Emily J. Fray, Fengting Wu, Francesco R. Simonetti, Carolin Zitzmann, Narmada Sambaturu, Carmen Molina-Paris, Alexandra M. Bender, Po-Ting Liu, John D. Ventura, Roger W. Wiseman, David H. O'Connor, Romas Gelezunas, Thomas Leitner, Ruy M. Ribeiro, Alan S. Perelson, Dan H. Barouch, Janet D. Siliciano, and Robert F. Siliciano

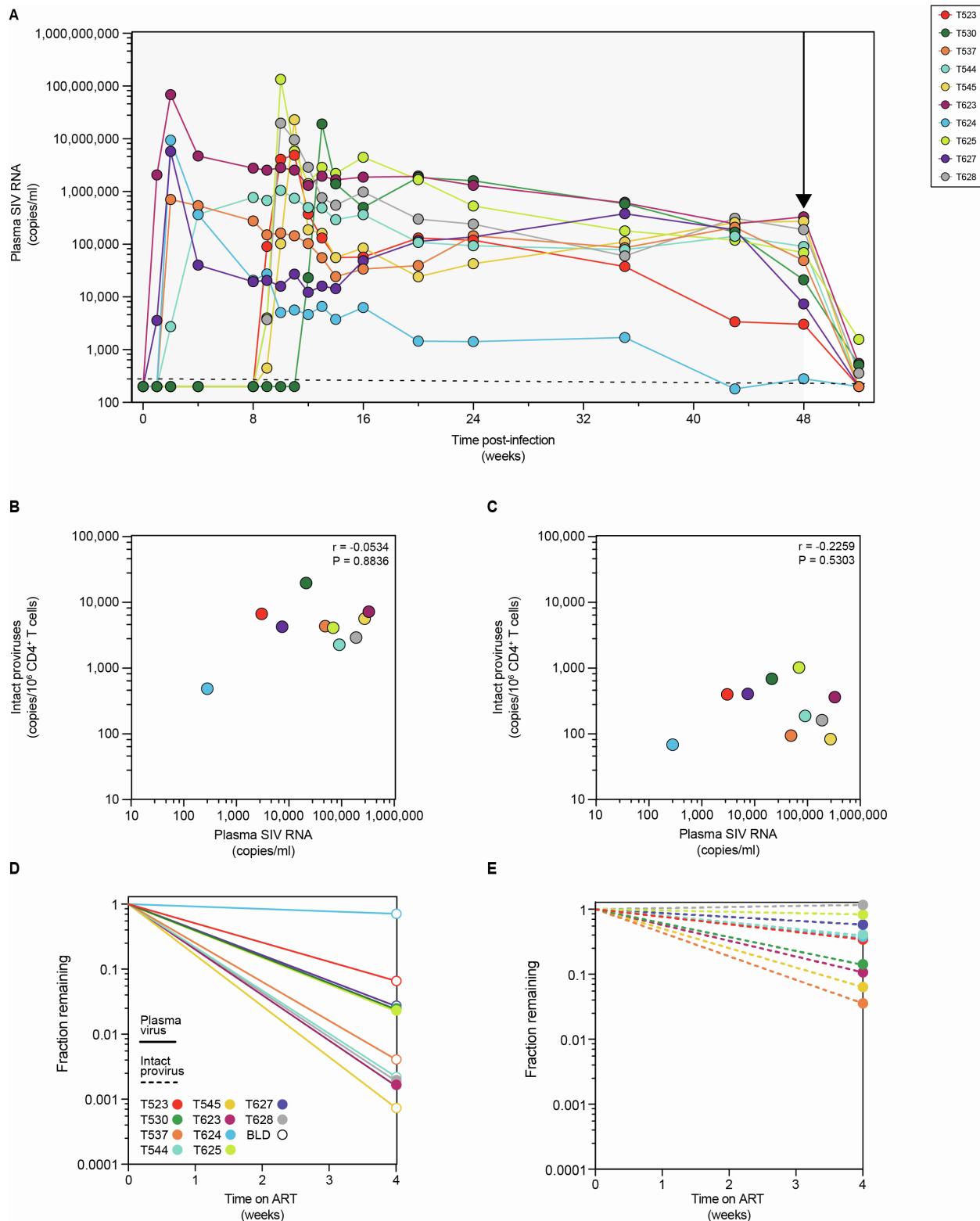


Figure S1. Dynamics of plasma virus and intact proviruses (related to Figure 1). (A) Pre-ART plasma viral loads for cohort 18-02. Animals were infected via repetitive intrarectal challenge with SIVmac₂₅₁ until plasma became positive for SIV RNA. ART was initiated after ~48 weeks (black arrow). The L.O.D. (200 copies/ml) is indicated by the horizontal dotted line. (B) Correlation between plasma SIV RNA at the time of ART initiation and intact proviruses quantified using the IPDA at ART initiation (t=0) or (C) week

92 on ART. Individual datapoints represent the number of intact proviruses measured in CD4⁺ T cell DNA samples using the IPDA (y-axis)¹, or plasma SIV RNA measured using *gag* qPCR (x-axis)². IPDA measurements represent the geometric mean of 3 replicates. Intact proviruses are corrected for shearing and *env*⁺2LTR circles. The correlation between the two variables was calculated using Pearson's coefficient. (D) Decay of plasma viremia and, (E) intact proviruses after 4 weeks of ART. Lines indicate fractional decrease in plasma virus (left, solid lines) or intact proviruses (right, dotted lines) between the baseline (week 0) and first on-ART measurement (week 4). For all but 4 animals, plasma virus was undetectable (less than 200 copies/ml) at week 4, as indicated by the open circles.

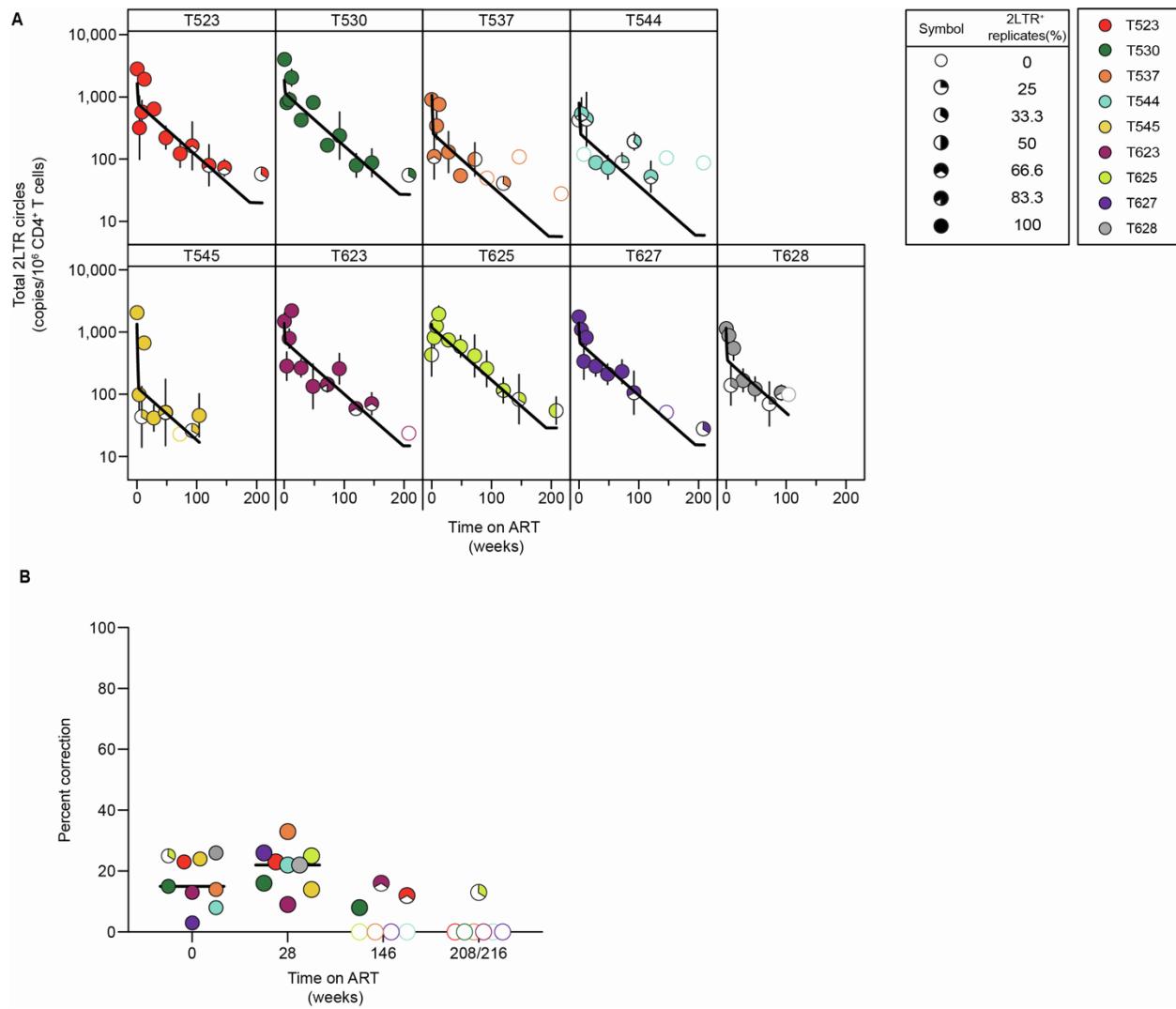


Figure S2. Decay of 2LTR circles after treatment initiation (related to Figure 2). (A) Decay of total 2LTR circles on ART. Data were fitted to a triphasic model (black line) as described in the Methods section. Vertical lines indicate the standard deviations. Degree of symbol filling indicates the number of replicates that were positive for 2LTR circles. Between 3 and 6 replicates were run per sample. Solid circles indicate all 3 replicates were positive for 2LTR circles, and open circles indicate that 0 replicates were positive. (B) Percent correction of IPDA data using *env*⁺2LTR events at weeks 0, 28, 146, or 208 (animals T523, T530, T544, T623, T625 and T627) or 216 (animal T537) on ART. Open circles indicate that no *env*⁺2LTR circles were detected, and therefore no correction was applied to the IPDA values. Horizontal black lines represent the mean. Animal T624 was not included in this analysis (see Methods).

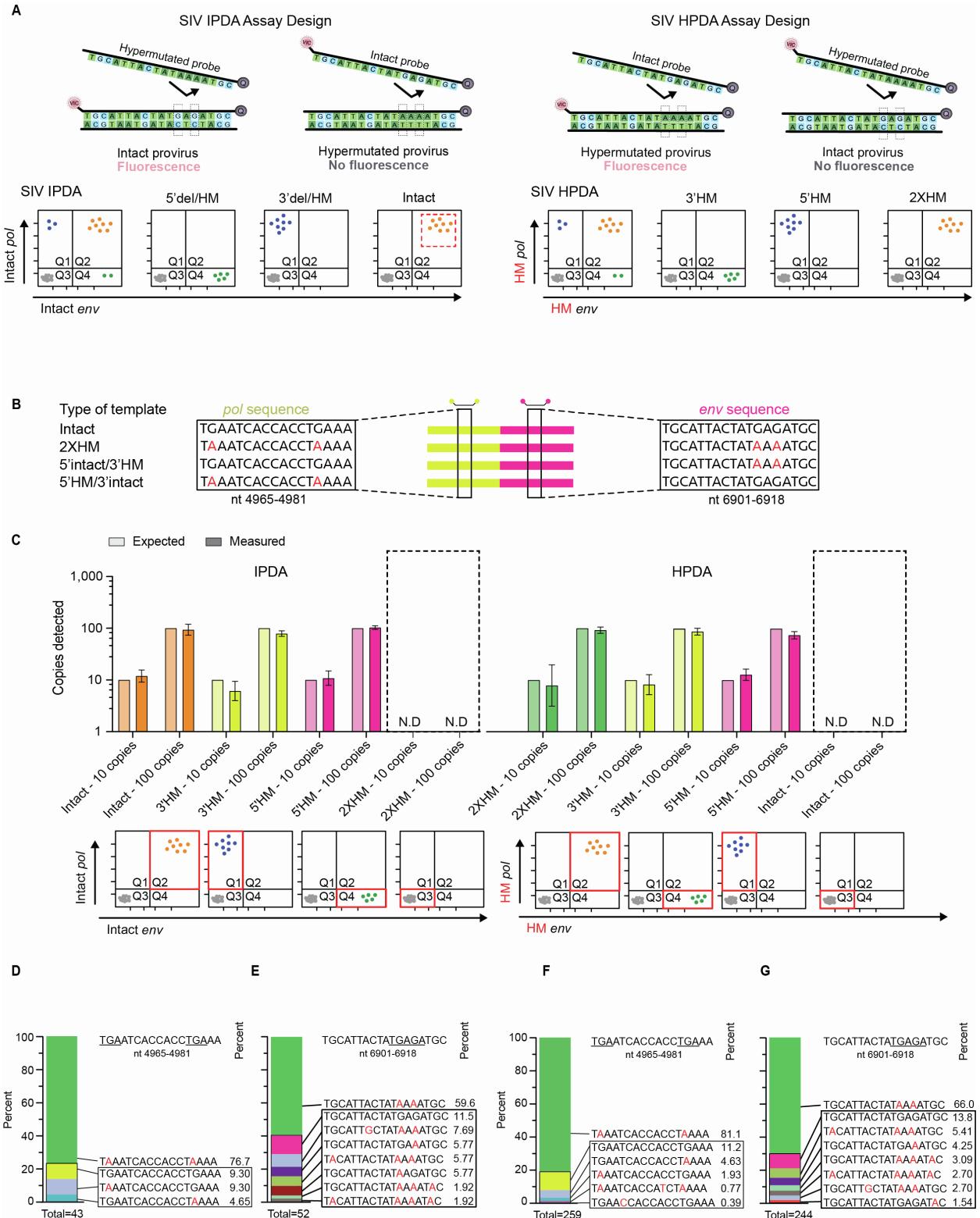


Figure S3. Design and validation of a ddPCR assay for hyperpermuted SIV proviruses (related to STAR Methods). (A) Comparison of assay design between the previously described Intact Proviral DNA Assay (IPDA)¹ and the Hyperpermuted Proviral DNA Assay (HPDA). The IPDA relies on labeled probes that recognize intact proviruses,

and unlabeled competition probes that exclude hypermutated proviruses from detection. Events are scored as positive for *pol* and/or *env* as described in Fig. 1C. The HPDA utilizes labeled hypermutated probes to detect proviruses with G→A mutation at key positions, and unlabeled probes matching intact sequences are used to exclude intact proviruses from detection. (B) Synthetic double-stranded DNA templates used as assay controls. The regions of the SIVmac₂₅₁ genome corresponding to probe sequences are indicated. (C) Results of control experiments using the IPDA or HPDA. 10 or 100 copies of each of the templates depicted in panel B were tested in quadruplicate with either the IPDA or HPDA primers and probes. Values represent the geometric mean calculated across 4 independent replicates, and bars indicate the standard deviation. The predicted quadrant in which positive events for each control template should be detected by either assay is indicated below, in red. N.D. indicates that no positive events were detected. (D) Summary of sequence polymorphisms (red) in the *pol* or (E) *env* probe binding site, identified in a dataset of previously published sequences collected from SIV-infected macaques¹. (F) Summary of sequence polymorphisms (red) in the *pol* or (G) *env* probe binding site identified in a second dataset of nFL hypermutated sequences from SIV-infected macaques³. Sequence variants with a frequency of <1% were not included in panel G and account for a small fraction of the total number of sequences (15 total, ~6%). Sites that are subject to APOBEC3G/F-mediated G→A hypermutation and therefore used to discriminate intact and hypermutated proviruses are underlined.

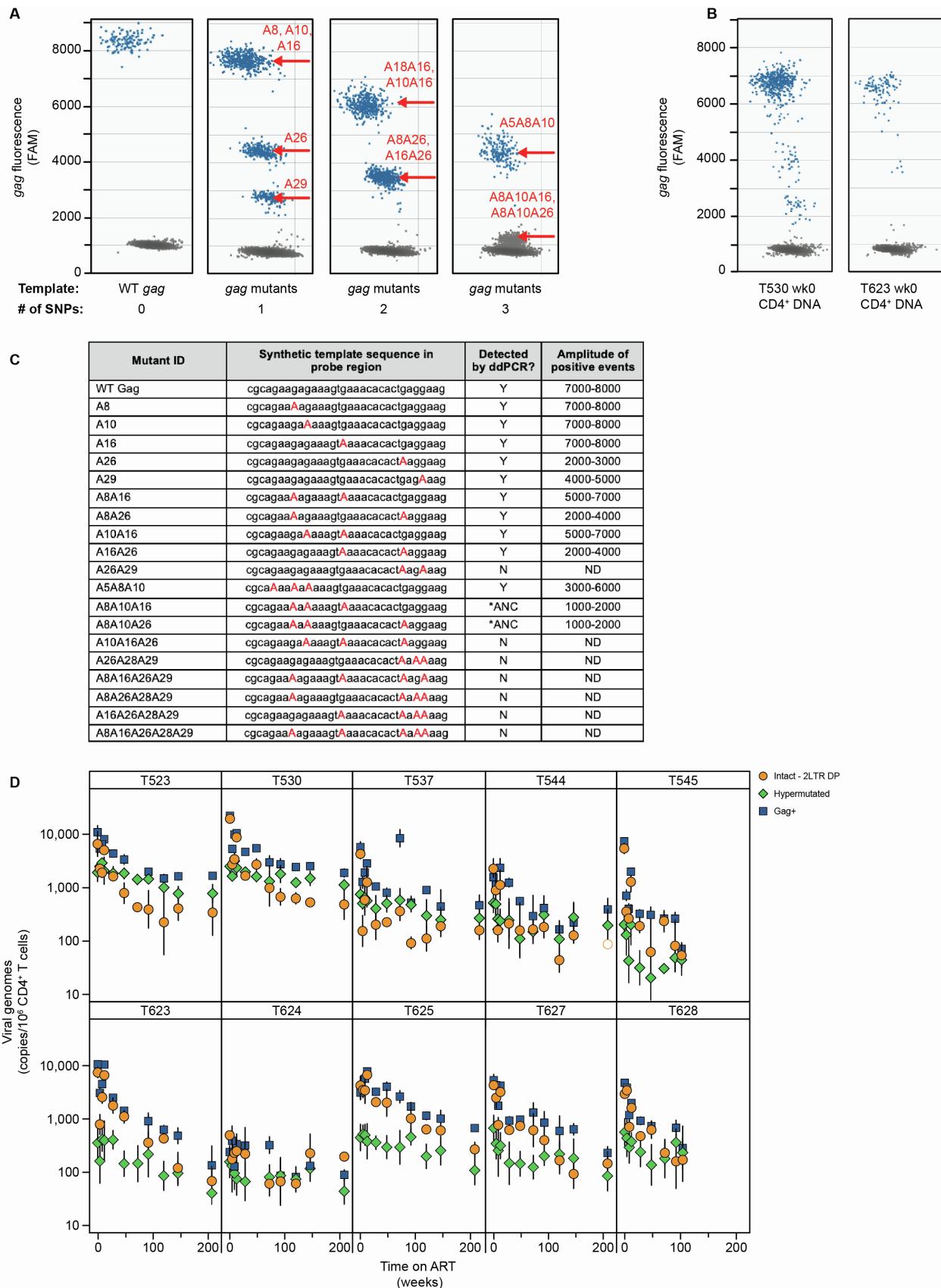


Figure S4. gag assay performance, detection of templates with G→A mutations and comparison to intact and hypermutated proviruses (related to Figure 4). (A) Dot plot images depicting representative *gag* ddPCR⁴ results using synthetic double-stranded DNA templates with 1-to-3 polymorphisms in the probe binding site. The images show the results from 3 independent replicates of each indicated template, with 250 copies of template loaded per reaction. The mutations that were tested and the corresponding mutant IDs/names are listed in the table below. (B) Representative dot plot images displaying data for 3 replicates of the *gag* assay when used on samples from SIV-infected macaques. An animal with a high abundance of hypermutated proviruses (T530) and one with a low abundance of hypermutated proviruses (T623) are depicted. (C) Table describing the *gag* polymorphisms tested using synthetic DNA templates. Mutant IDs indicate the position(s) in the probe binding site where G→A mutations (red, capital letters) were introduced. Each mutant template was scored according to whether it is detectable by ddPCR with the *gag* probe, and the relative range of the amplitude of positive events for that template. ANC = “above negative cluster,” which indicates that positive events are detected above the negative droplet cluster but there is insufficient separation to accurately gate or quantify. ND = “not detected,” which indicates that no difference in fluorescence amplitude was observed relative to negative controls. (D) Comparison of intact, hypermutated, and *gag*⁺ SIV DNA decay. Decay of intact (orange circles), hypermutated (green diamonds) and *gag*⁺ (navy squares) SIV DNA during ART. Intact provirus values are corrected for *env*⁺2LTR circles and shearing; and hypermutated provirus values are corrected for shearing. Vertical lines represent the standard deviations.

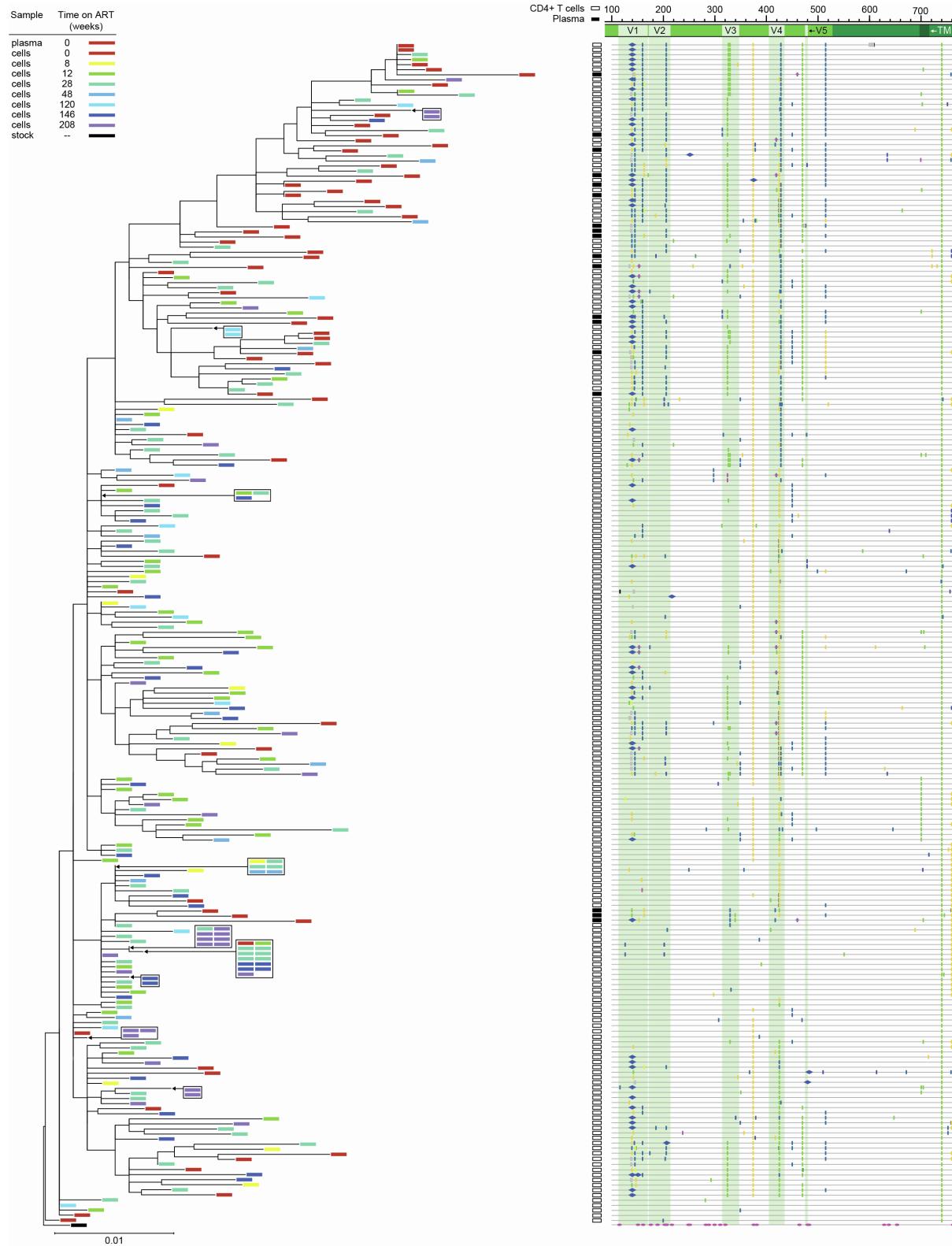


Figure S5. Sequencing analysis of *env* diversity in plasma virus and proviral DNA from CD4⁺ T cells for T523 (related to Figure 6). Phylogenetic tree and highlighter plot of *env* sequences collected from either plasma or CD4⁺ T cell DNA from animal T523. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

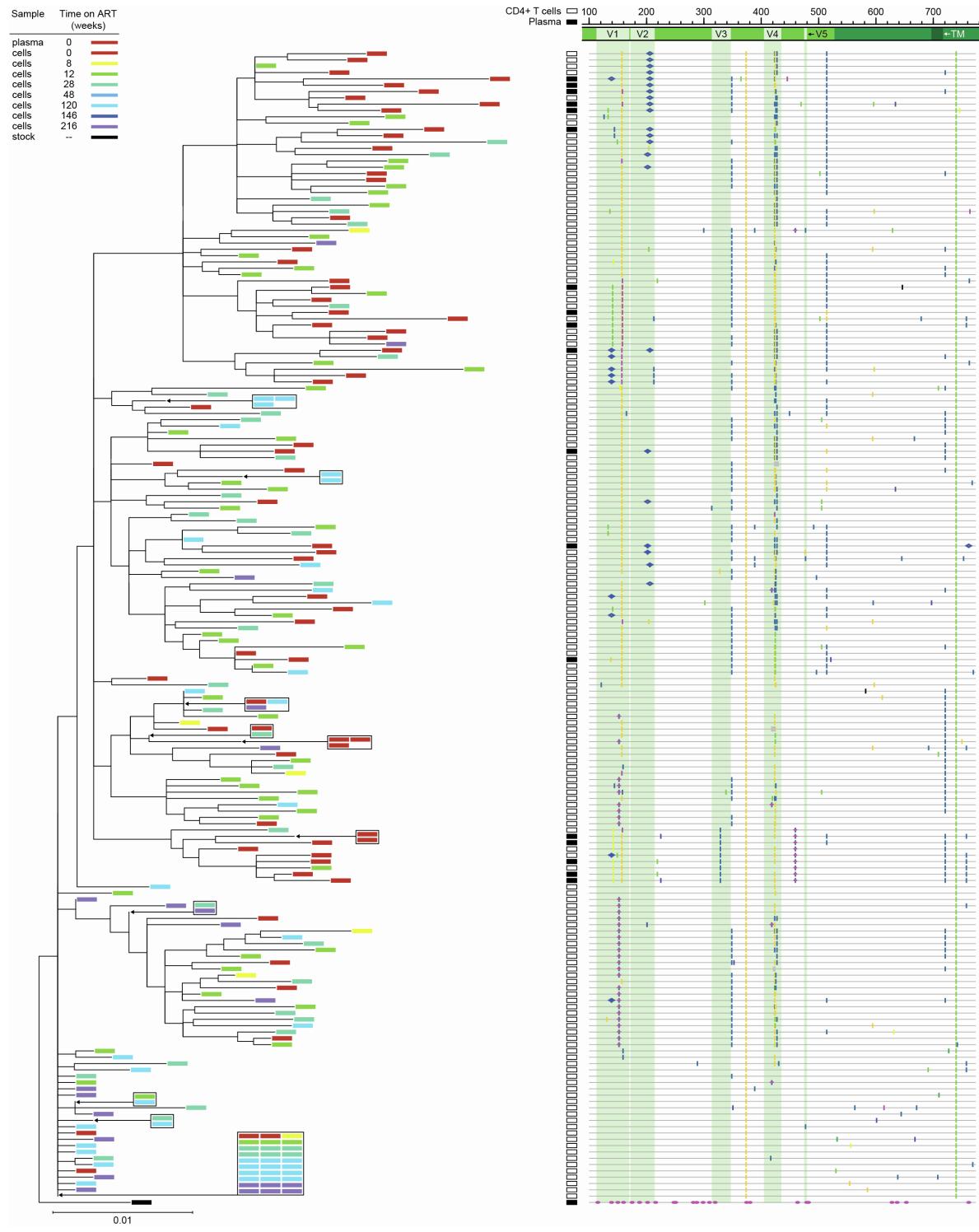


Figure S6. Sequencing analysis of *env* diversity in plasma virus and proviral DNA from CD4⁺ T cells for T537 (related to Figure 6). Phylogenetic tree and highlighter plot of *env* sequences collected from either plasma or CD4⁺ T cell DNA from animal T537. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

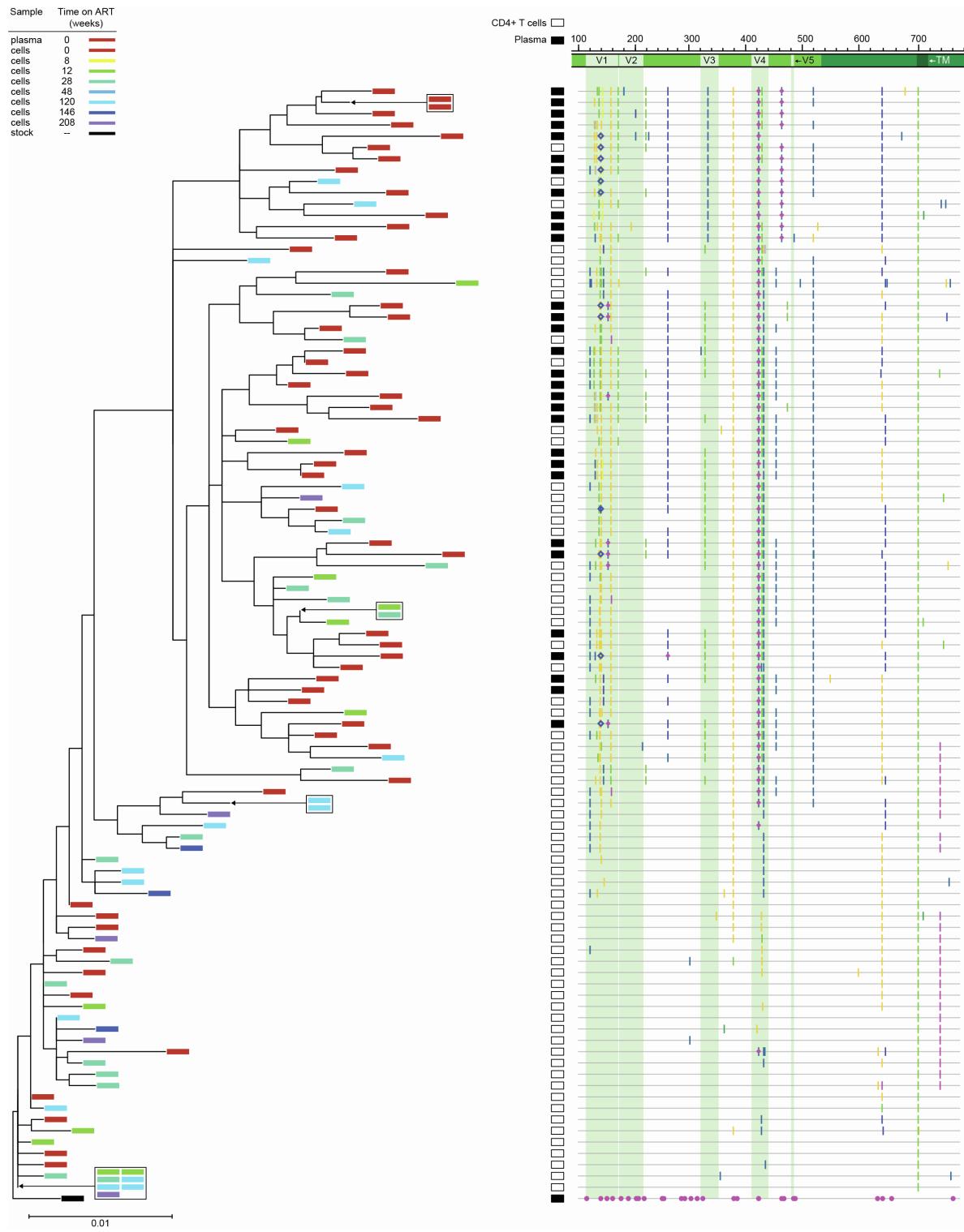


Figure S7. Sequencing analysis of *env* diversity in plasma virus and proviral DNA from CD4⁺ T cells for T544 (related to Figure 6). Phylogenetic tree and highlighter plot of *env* sequences collected from either plasma or CD4⁺ T cell DNA from animal T544. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

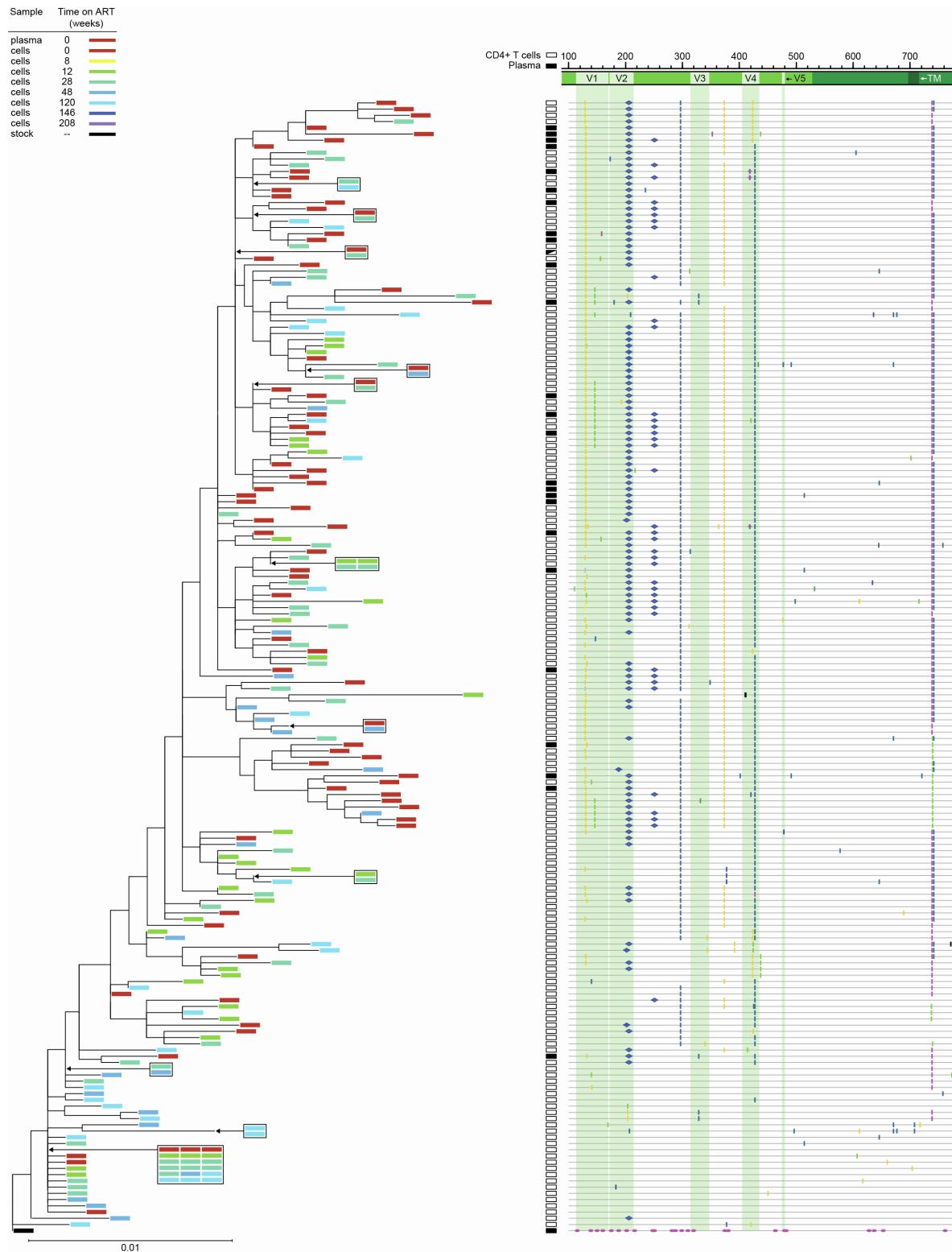


Figure S8. Sequencing analysis of env diversity in plasma virus and proviral DNA from CD4⁺ T cells for T545 (related to Figure 6). Phylogenetic tree and highlighter plot of env sequences collected from either plasma or CD4⁺ T cell DNA from animal T545. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

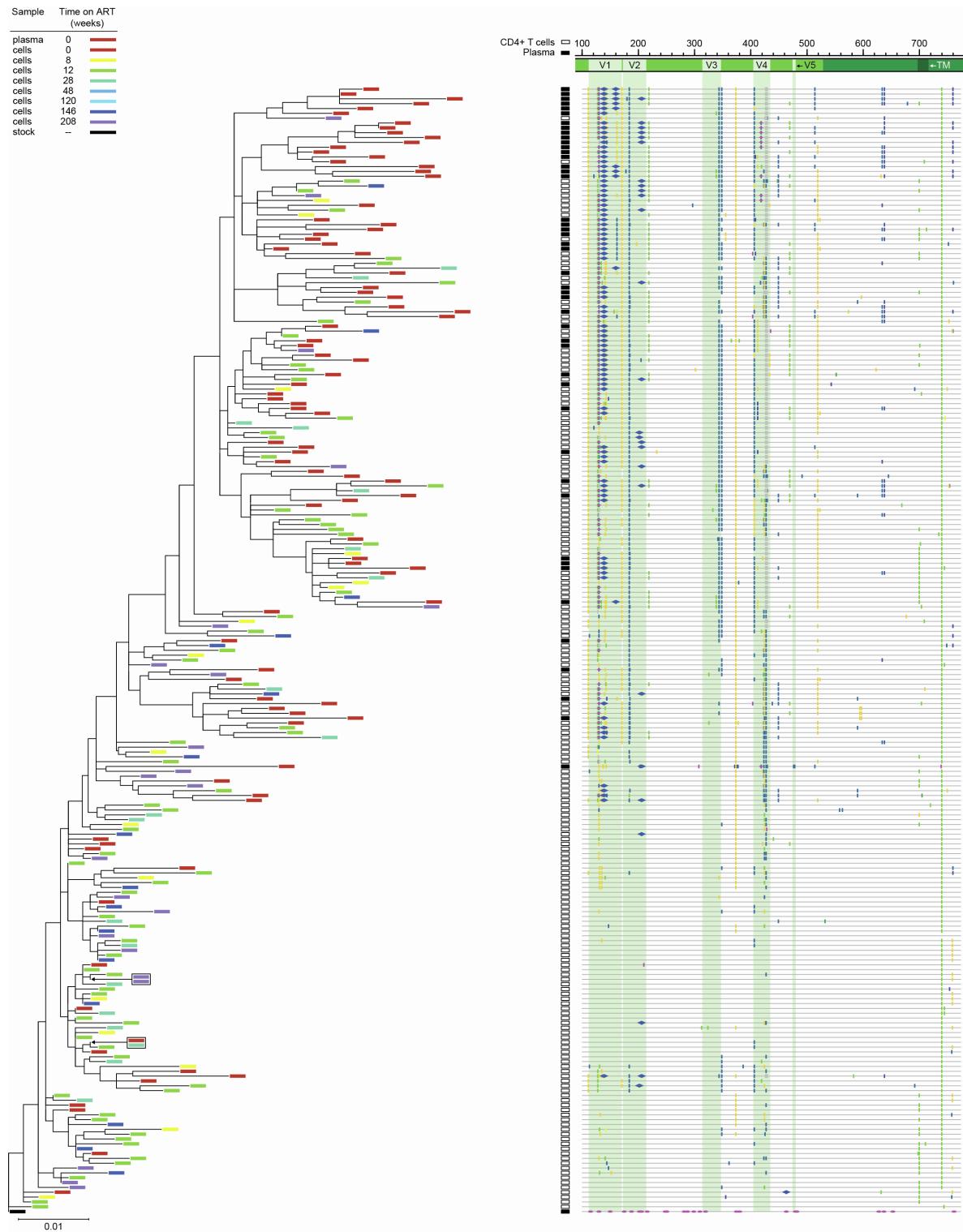


Figure S9. Sequencing analysis of *env* diversity in plasma virus and proviral DNA from CD4⁺ T cells for T623 (related to Figure 6). Phylogenetic tree and highlighter plot of *env* sequences collected from either plasma or CD4⁺ T cell DNA from animal T623. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

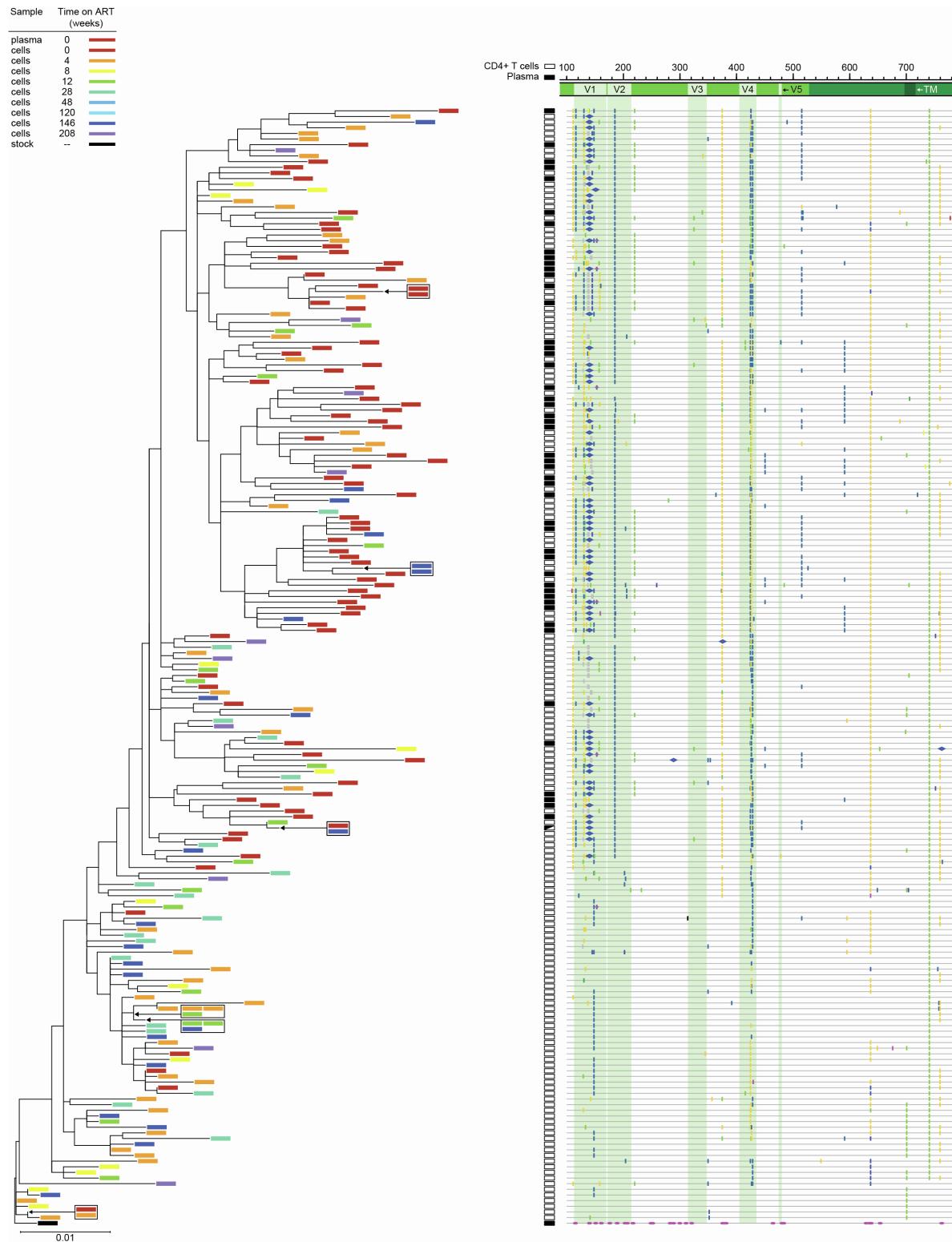


Figure S10. Sequencing analysis of env diversity in plasma virus and proviral DNA from CD4⁺ T cells for T625 (related to Figure 6). Phylogenetic tree and highlighter plot of env sequences collected from either plasma or CD4⁺ T cell DNA from animal T625. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

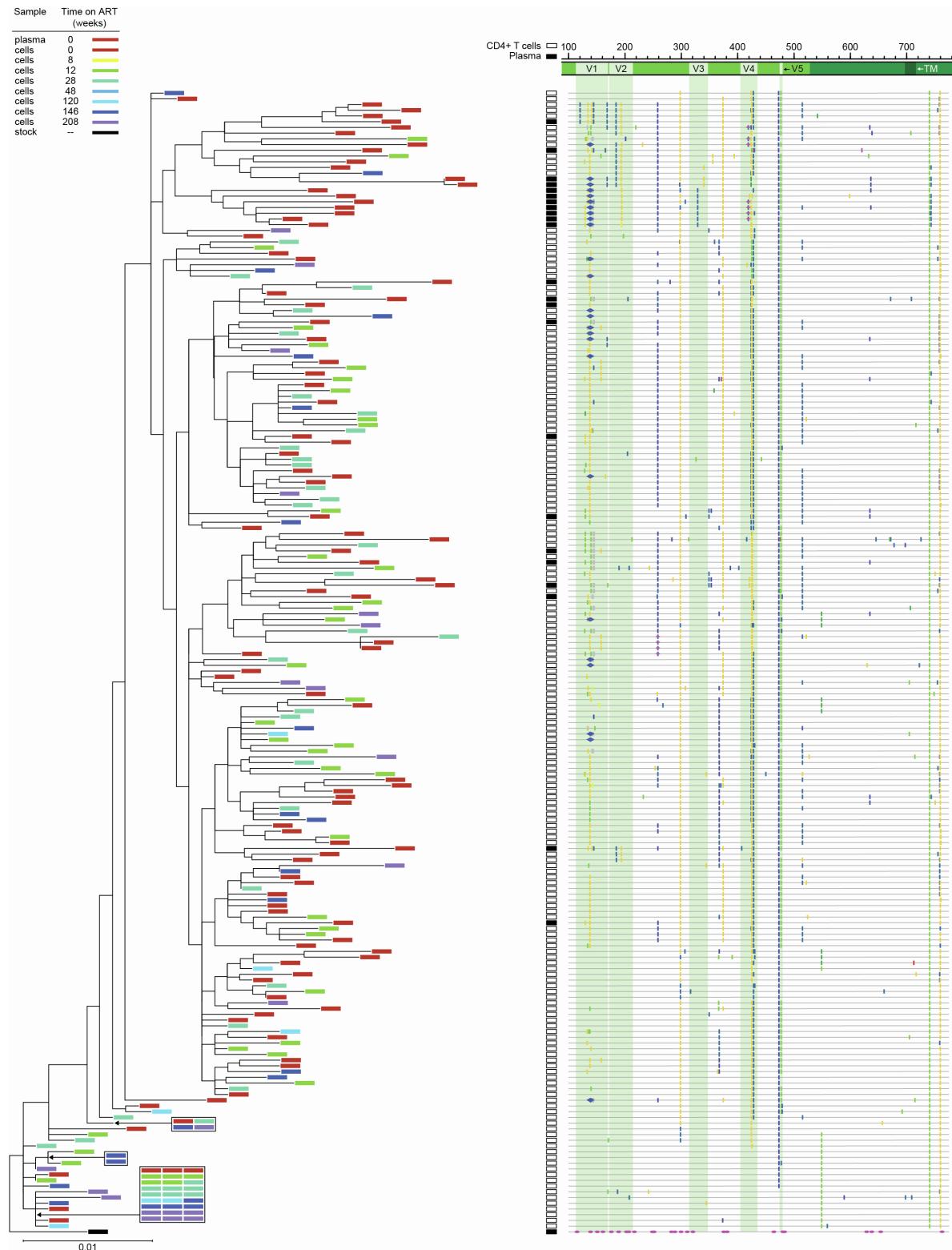


Figure S11. Sequencing analysis of *env* diversity in plasma virus and proviral DNA from CD4⁺ T cells for T627 (related to Figure 6). Phylogenetic tree and highlighter plot of *env* sequences collected from either plasma or CD4⁺ T cell DNA from animal T627. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

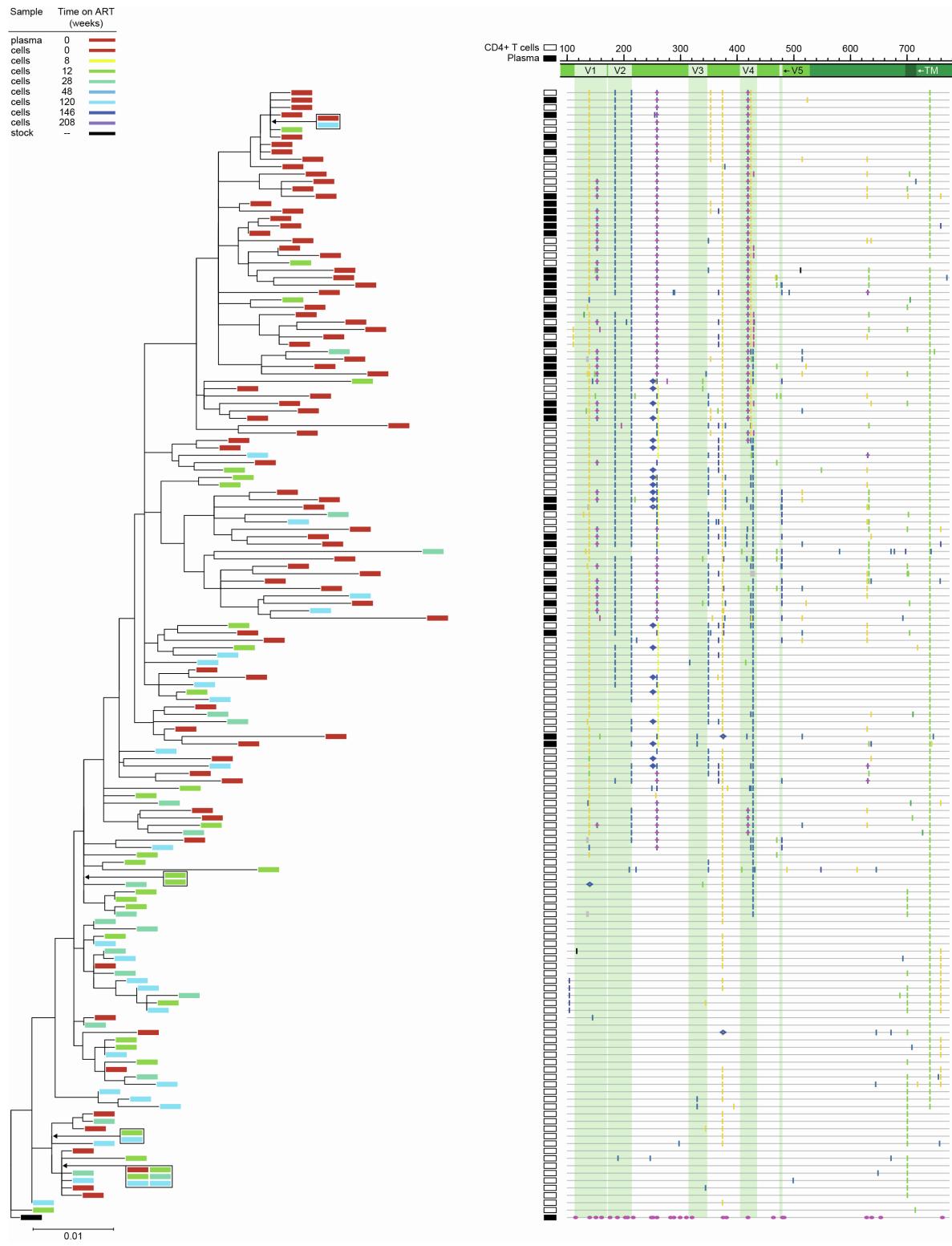


Figure S12. Sequencing analysis of *env* diversity in plasma virus and proviral DNA from CD4⁺ T cells for T628 (related to Figure 6). Phylogenetic tree and highlighter plot of *env* sequences collected from either plasma or CD4⁺ T cell DNA from animal T628. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

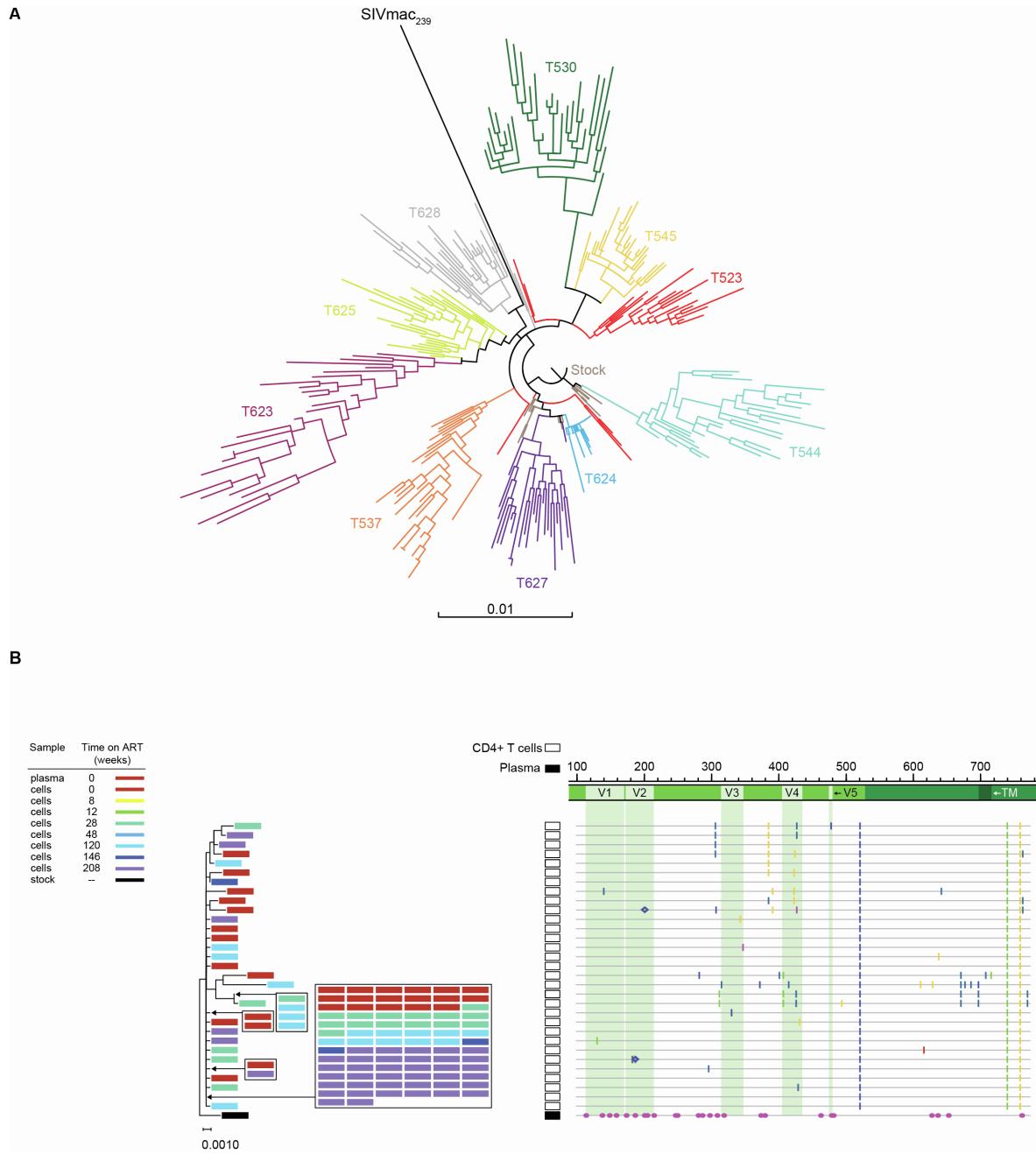


Figure S13. Viral replication during untreated SIV infection leads to distinct patterns of intra-individual diversification (related to Figure 6). (A) Composite tree of env sequences from animals in cohort 18-02. The tree is rooted on a consensus sequence of the stock⁵. To capture the maximum divergence observed in each animal, tree was made with 25-30 randomly chosen time 0 plasma and cellular sequences from each animal. For animal T624, only cellular sequences were available. Sequence analysis includes 2054 nt. (B) Sequencing analysis of env diversity in proviral DNA from CD4⁺ T cells for T624. Phylogenetic tree and highlighter plot of env sequences collected from CD4⁺ T cell DNA from animal T624. The tree is rooted on a consensus sequence of the stock⁵. Sequence analysis includes 2054 nt.

Table S1. Decay of SIV DNA forms following initiation of ART[†] (related to Figures 1-4).

<i>y</i>	<i>y</i> ₀ (log10)	<i>A</i>	<i>b</i> ₁ (/week)	<i>t</i> _{1/2} for initial phase (days)	<i>b</i> ₂ (/week)	<i>t</i> _{1/2} for second phase (weeks)	<i>T</i> _s (weeks)
Intact proviruses [†]	3.60 (3.36-3.86)	0.638 (0.50-0.82)	1.45 (1.08-1.95)	3.3 (2.5-4.5)	0.020 (0.016-0.024)	35 (28-43)	119 (86-153)
HM [#] – biphasic	2.68 (2.46-2.92)	0.694 (0.57-0.78)	0.117 (0.09-0.154)	42 (31-55)	0 (0.0-10 ⁻⁴)	No decay	-
HM [#] – monophasic	2.68 (2.46-2.92)	0	-	-	0.0058 (0.004-0.008)	120 (84-172)	-
<i>gag</i> ⁺ DNA [§]	3.64 (3.34-3.96)	0.420 (0.19-0.69)	0.743 (0.653-0.816)	6.5 (5.9-7.4)	0.012 (0.011-0.014)	56 (51-61)	149 (118-187)
Total 2LTR [¶] circles	3.09 (2.94-3.23)	0.521 (0.35-0.69)	1.42 (NA) [#]	3.4 (NA) [#]	0.019 (0.016-0.023)	36 (30-43)	195 (157-233)
<i>env</i> ⁺ 2LTR [¶] circles	2.92 (2.74-3.11)	0.501 (0.18-0.82)	0.948 (0.61-1.00)	5.1 (4.9-8.0)	0.020 (0.017-0.025)	34 (28-41)	-
<i>env</i> 2LTR [¶] circles	2.34 (2.13-2.57)	0.561 (0.19-0.70)	0.516 (0.33-0.70)	9.4 (6.9-14.9)	0.013 (0.010-0.016)	53 (43-66)	-

*We used a non-linear mixed effect approach to fit multiple models for the decay of SIV DNA species (see Methods). *y* is the variable of interest, *y*₀ is the baseline value, *A* is the fraction of *y* that decays in the first phase with decay rate *b*₁ and (1 – *A*) is the fraction of *y* that decays in the second phase with decay rate *b*₂. After this potential biphasic decay lasting until time *T*_s, there is a third phase with decay rate *b*₃, which would be smaller than the two initial phases. The table shows the best fit population parameter estimates and 95% confidence intervals.

[†]Intact proviruses were quantified using the double-positive population from the SIV IPDA and corrected for shearing and *env*⁺2LTR circles.

[#]HM proviruses were quantified by adding the positive events from quadrants 1, 2, and 4 from the SIV HPDA and corrected for shearing.

[§]*gag*⁺ DNA was quantified using the positive droplets from the *gag* simplex assay.

[¶]2LTR circles were quantified using the *env*-2LTRc duplex assay. Total 2LTR circles were calculated by adding the positive events from quadrants 1 (*env*) and 2 (*env*⁺).

[#]This confidence interval could not be estimated.

Table S2. Decay of SIV DNA forms following initiation of ART, excluding week 12*
(related to Figures 1-4).

<i>y</i>	<i>y</i> ₀ (log10)	<i>A</i>	<i>b</i> ₁ (/week)	<i>t</i> _{1/2} for initial phase (days)	<i>b</i> ₂ (/week)	<i>t</i> _{1/2} for second phase (weeks)	<i>T</i> _s (weeks)
Intact proviruses [†]	3.61	0.784	1.56	3.1	0.015	46	136
HM [‡] – biphasic	2.70	0.718	0.19	26	0	No decay	-
HM [‡] – monophasic	2.70	0	-	-	0.0057	122	-
<i>gag</i> ⁺ DNA [§]	3.64	0.554	0.69	7.0	0.0096	72	171
Total 2LTR [¶] circles	3.09	0.660	1.70	2.9	0.016	43	190
<i>env</i> ⁺ 2LTR [¶] circles	2.92	0.766	1.06	4.6	0.016	43	-
<i>env</i> 2LTR [¶] circles	2.36	0.553	0.61	8.0	0.012	58	-

*A non-linear mixed effect approach was used to fit multiple models for the decay of intact proviruses (see Methods). See Table S1 for description of variables. In the results present in this table, we excluded the week 12 datapoint from the fits, because it showed an unexpected increase in SIV DNA (see main text). Notice that most parameters change by a small amount, and in particular, the half-lives change by less than 30%.

[†]Intact proviruses were quantified using the double-positive population from the SIV IPDA and corrected for shearing and *env*⁺2LTR circles.

[‡]HM proviruses were quantified by adding the positive events from quadrants 1, 2, and 4 from the SIV HPDA and corrected for shearing.

[§]*gag*⁺ DNA was quantified using the positive droplets from the *gag* simplex assay.

[¶]2LTR circles were quantified using the *env*-2LTRc duplex assay. Total 2LTR circles were calculated by adding the positive events from quadrants 1 (*env*⁺) and 2 (*env*⁺).

Table S3. Individual parameter estimates for the decay of intact proviruses following initiation of ART* (related to Figure 2).

Animal ID	Baseline plasma SIV RNA (copies/ml)	Intact proviruses						
		y_0 (log10)	A	First phase		Second phase		
				b_1 (/week)	$t_{1/2}$ (days)	b_2 (/week)	$t_{1/2}$ (weeks)	
T523	3030	3.8	0.61	1.453	3.34	0.0200	35	108
T530	21190	4.1	0.66	1.454	3.34	0.0199	35	111
T537	49037	3.5	0.79	1.460	3.32	0.0195	36	92
T544	90354	3.3	0.73	1.451	3.34	0.0199	35	121
T545	273304	3.6	0.89	1.455	3.33	0.0199	35	119
T623	330056	3.8	0.62	1.455	3.33	0.0200	35	164
T624	281	2.9	0.64	1.454	3.34	0.0197	35	67
T625	69210	3.9	0.52	1.453	3.34	0.0196	35	114
T627	7404	3.7	0.61	1.452	3.34	0.0198	35	138
T628	189002	3.5	0.62	1.450	3.35	0.0203	34	119
Median	59124	3.7	0.63	1.453	3.34	0.0199	35	116
Mean	30508 [†]	3.6	0.67	1.454	3.34	0.0199	35	115
SD		0.3	0.11	0.003	0.01	0.0002	0.4	25

*A non-linear mixed effect approach was used to fit multiple models for the decay of intact proviruses (see Methods). This table shows best fit parameter estimates for individual animals. See Table S1 for description of variables.

[†]Geometric mean for all animals.

Table S4. Individual parameter estimates for decay of total 2LTR circles following initiation of ART* (related to Figure 2).

Animal ID	Baseline plasma SIV RNA (copies/ml)	Total 2LTR circles						
		y_0 (log10)	A	First phase		Second phase		Stable phase
				b_1 (/week)	$t_{1/2}$ (days)	b_2 (/week)	$t_{1/2}$ (weeks)	T_s (weeks)
T523	3030	3.2	0.52	1.43	3.40	0.0193	36	189
T530	21190	3.3	0.37	1.42	3.42	0.0193	36	193
T537	49037	3.0	0.76	1.44	3.37	0.0194	36	196
T544	90354	2.9	0.67	1.39	3.48	0.0193	36	195
T545	273304	3.1	0.90	1.44	3.38	0.0193	36	195
T623	330056	3.1	0.49	1.43	3.39	0.0193	36	200
T624	281	ND [†]						
T625	69210	3.1	0.10	1.42	3.42	0.0193	36	192
T627	7404	3.1	0.52	1.41	3.45	0.0194	36	195
T628	189002	3.1	0.69	1.30	3.73	0.0194	36	195
Median	59124	3.1	0.52	1.42	3.42	0.0193	36	195
Mean	30508 [‡]	3.1	0.56	1.41	3.45	0.0193	36	194
SD		0.1	0.23	0.04	0.11	0.00002	0.04	3

*A non-linear mixed effect approach was used to fit multiple models for the decay of 2LTR circles (see Methods). This table shows best fit parameter estimates for individual animals. See Table S1 for description of variables.

[†]ND. Not done. 2LTR decay for T624 could not be fit, likely due to values below the limit of detection at multiple time points. This animal had relatively low viral loads (Fig. S1A), and was excluded from this analysis.

[‡]Geometric mean for all animals.

Table S5. Individual parameter estimates for decay of hypermutated proviruses following initiation of ART* (related to Figure 3).

Animal ID	Baseline plasma SIV RNA (copies/ml)	Total HM proviruses					
		y_0 (log10)	A	First phase		Second phase	
				b_1 (/week)	$t_{1/2}$ (days)	b_2 (/week)	$t_{1/2}$ (weeks)
T544	90354	2.6	0.64	0.118	41	0	-
T545	273304	2.2	0.80	0.115	42	0	-
T624	281	2.2	0.63	0.120	40	0	-
T627	7404	2.7	0.73	0.117	41	0	-
T628	189002	2.7	0.65	0.117	42	0	-
Median	90354	2.6	0.65	0.117	41	0	-
Mean	24972 [†]	2.5	0.69	0.117	41	0	-
SD		0.3	0.07	0.002	0.7	0	-
T523	3030	3.4	0	-	-	0.0059	118
T530	21190	3.4	0	-	-	0.0046	151
T537	49037	2.8	0	-	-	0.0051	136
T623	330056	2.5	0	-	-	0.0089	78
T625	69210	2.6	0	-	-	0.0059	117
Median	49037	2.8	0	-	-	0.0059	118
Mean	37271 [†]	2.9	0	-	-	0.0061	120
SD		0.4	-	-	-	0.0017	28

*A non-linear mixed effect approach was used to fit multiple models for the hypermutated proviruses (see Methods). The best fit was obtained for a hybrid model in which some macaques (T544, T545, T624, T627 and T628) had a biphasic decay, with a second phase decay rate that was essentially 0, and another group (T523, T530, T537, T623 and T625) had a single phase of decay. This table shows best fit parameter estimates for individual animals. See Table S1 for description of variables.

[†]Geometric mean for this group.

Table S6. Individual parameter estimates for decay of gag⁺ DNA following initiation of ART* (related to Figure 4).

Animal ID	Baseline plasma SIV RNA (copies/ml)	gag ⁺ DNA						
		y ₀ (log10)	A	First phase		Second phase		
				b ₁ (/week)	t _{1/2} (days)	b ₂ (/week)	t _{1/2} (weeks)	
T523	3030	4.0	0.36	0.74	6.5	0.0120	58	127
T530	21190	4.3	0.52	0.74	6.5	0.0115	60	129
T537	49037	3.7	0.52	0.74	6.5	0.0103	67	154
T544	90354	3.2	0.45	0.74	6.5	0.0117	59	120
T545	273304	3.7	0.87	0.75	6.5	0.0158	44	149
T623	330056	4.0	0.48	0.74	6.5	0.0167	41	201
T624	281	2.6	0.25	0.74	6.5	0.0107	65	125
T625	69210	3.8	0.16	0.74	6.5	0.0117	59	163
T627	7404	3.7	0.51	0.74	6.5	0.0123	57	181
T628	189002	3.7	0.61	0.74	6.6	0.0146	48	149
Median	59124	3.7	0.50	0.74	6.5	0.0119	58	149
Mean	30508 [†]	3.7	0.47	0.74	6.5	0.0127	56	150
SD		0.5	0.19	0.002	0.01	0.0022	9	26

*A non-linear mixed effect approach was used to fit multiple models for the decay of gag⁺ DNA (see Methods). This table shows best fit parameter estimates for individual animals. See Table S1 for description of variables.

[†]Geometric mean for all animals.

Table S7. Number of sequences analyzed per animal for average pairwise distance calculations (related to Figures 5 & 7).

Animal ID	Parameter	Plasma wk0	CD4 ⁺ DNA wk0-12	CD4 ⁺ DNA wk28-48	CD4 ⁺ DNA wk104-216
T523	n	18	96	84	75
	Mean APD	0.007278	0.005665	0.004878	0.004522
	Std deviation	0.001315	0.002258	0.002335	0.001933
T530	n	89	99	85	103
	Mean APD	0.01355	0.01025	0.00868	0.007032
	Std deviation	0.001317	0.003961	0.004001	0.003782
T537	n	20	74	40	62
	Mean APD	0.008728	0.006213	0.00568	0.004624
	Std deviation	0.0008068	0.001477	0.001655	0.001784
T544	n	31	35	18	24
	Mean APD	0.01002	0.006148	0.005749	0.005236
	Std deviation	0.0009446	0.002921	0.002453	0.002535
T545	n	27	125	59	28
	Mean APD	0.006739	0.005054	0.004327	0.003893
	Std deviation	0.001236	0.001944	0.002107	0.002008
T623	n	48	88	19	38
	Mean APD	0.01256	0.007766	0.006607	0.006386
	Std deviation	0.001828	0.003449	0.003774	0.003612
T624	n	na	30	18	64
	Mean APD	na	0.003709	0.003422	0.003427
	Std deviation	na	0.0008112	0.0007534	0.0005286
T625	n	49	104	17	35
	Mean APD	0.008224	0.005246	0.004119	0.005227
	Std deviation	0.001037	0.002313	0.001483	0.002232
T627	n	24	116	47	53
	Mean APD	0.0101	0.007767	0.006914	0.006073
	Std deviation	0.001312	0.002314	0.002609	0.002642
T628	n	36	76	18	28
	Mean APD	0.008602	0.005603	0.004178	0.004414
	Std deviation	0.001442	0.002539	0.00238	0.002215

Table S8. Droplet digital PCR assay primers and probes (related to STAR Methods).

Amplicon	Primer/Probe Name	Sequence	Fluoro-phore	Quencher
SIVmac ₂₅₁ IPDA <i>pol</i>	polF	GCA GGG ATA GAG CAC ACC TTT G	N/A	N/A
	polR	CTA TGG TTT CTA CTG AAT TTG CTT GTT C	N/A	N/A
	pol intact probe	TTT CAG GTG GTG ATT CA	FAM	MGBNFQ
	pol HM probe	TAG GTG GTG ATT TAT T	N/A	MGBNFQ
SIVmac ₂₅₁ IPDA <i>env</i>	envF	CCT CAA TAA AGC CTT GTG TAA AAT TAT C	N/A	N/A
	envR	GTT GTT GAT GAT TTT GTC AAT CCC	N/A	N/A
	env intact probe	TGC ATT ACT ATG AGA TGC	VIC	MGBNFQ
	env HM probe	TGC ATT ACT ATA AAA TGC	N/A	MGBNFQ
SIVmac ₂₅₁ G7 IPDA <i>env</i>	envF	CCT CAA TAA AGC CTT GTG TAA AAT TAT C	N/A	N/A
	envR	GTT GTT GAT GAT TTT GTC AAT CCC	N/A	N/A
	env intact G7 probe	TGC ATT GCT ATG AGA TGC	VIC	MGBNFQ
	env HM G7 probe	TGC ATT GCT ATA AAA TGC	N/A	MGBNFQ
	env intact probe	TGC ATT ACT ATG AGA TGC	VIC	MGBNFQ
	env HM probe	TGC ATT ACT ATA AAA TGC	N/A	MGBNFQ
SIVmac ₂₅₁ HPDA <i>pol</i>	polF	GCA GGG ATA GAG CAC ACC TTT G	N/A	N/A
	polR	CTA TGG TTT CTA CTG AAT TTG CTT GTT C	N/A	N/A
	pol intact probe	TTT CAG GTG GTG ATT CA	N/A	MGBNFQ
	pol HM probe	TAG GTG GTG ATT TAT T	FAM	MGBNFQ
SIVmac ₂₅₁ HPDA <i>env</i>	envF	CCT CAA TAA AGC CTT GTG TAA AAT TAT C	N/A	N/A
	envR	GTT GTT GAT GAT TTT GTC AAT CCC	N/A	N/A
	env intact probe	TGC ATT ACT ATG AGA TGC	N/A	MGBNFQ
	env HM probe	TGC ATT ACT ATA AAA TGC	VIC	MGBNFQ
SIVmac ₂₅₁ 2LTRc	2LTRc F	CGC CTG GTC AAC TCG GTA CTC	N/A	N/A
	2LTRc R	GGT ATG ATG CCT TCT TCC TTT TCT AAG	N/A	N/A
	2LTRc probe	CCC TGG TCT GTT AGG ACC CTT TCT GCT TTG	FAM	MGBNFQ
SIV <i>gag</i>	gagF	GTC TGC GTC ATC TGG TGC ATT C	N/A	N/A
	gagR	CAC TAG GTG TCT CTG CAC TAT CTG TTT TG	N/A	N/A
	gag probe	CTT CCT CAG TGT GTT TCA CTT TCT CTT CTG CG	FAM	ZEN/IABkFQ
Rhesus RPP30-1	RPP30-1F	AGG ATG CTC CGG GAG TAT GTA	N/A	N/A
	RPP30-1R	CCT GCT TGT CAC CTA TAT AAC AT	N/A	N/A
	RPP30-1 probe	TCA AGC TGG GAG ACG GAA GAG TCA GT	FAM	ZEN/IABkFQ
Rhesus RPP30-2	RPP30-2F	ACA GAC TCA CAC AAT TTA GG	N/A	N/A
	RPP30-2R	ACA TTC ATG CCA CTG CAC TC	N/A	N/A
	RPP30-2 probe	ACA GGG TCT CAC TTT GTT GTC CA	HEX	ZEN/IABkFQ

Supplemental References List

1. Bender, A.M., Simonetti, F.R., Kumar, M.R., Fray, E.J., Bruner, K.M., Timmons, A.E., Tai, K.Y., Jenike, K.M., Antar, A.A.R., Liu, P., *et al.* (2019). The landscape of persistent viral genomes in ART-treated SIV, SHIV, and HIV-2 infections. *Cell Host Microbe.* 26, 73-85.e4.
2. Barouch, D.H., Ghneim, K., Bosche, W.J., Li, Y., Berkemeier, B., Hull, M., Bhattacharyya, S., Cameron, M., Liu, J., Smith, K., *et al.* (2016). Rapid inflammasome activation following mucosal SIV infection of rhesus monkeys. *Cell.* 165, 656-667.
3. Liu, P., Keele, B.F., Abbink, P., Mercado, N.B., Liu, J., Bondzie, E.A., Chandrashekhar, A., Borducchi, E.N., Hesselgesser, J., Mish, M., *et al.* (2020). Origin of rebound virus in chronically SIV-infected rhesus monkeys following treatment discontinuation. *Nat Commun.* 11, 5412.
4. Cline, A.N., Bess, J.W., Piatak, M., and Lifson, J.D. (2005). Highly sensitive SIV plasma viral load assay: Practical considerations, realistic performance expectations, and application to reverse engineering of vaccines for AIDS. *J Med Primatol.* 34, 303-312.
5. Del Prete, G.Q., Scarlotta, M., Newman, L., Reid, C., Parodi, L.M., Roser, J.D., Oswald, K., Marx, P.A., Miller, C.J., Desrosiers, R.C., *et al.* (2013). Comparative characterization of transfection- and infection-derived simian immunodeficiency virus challenge stocks for in vivo nonhuman primate studies. *J Virol.* 87, 4584-4595.