A majority of HIV persistence during antiretroviral therapy is due to infected cell proliferation

Daniel B. Reeves[1], Elizabeth R. Duke[1,2], Thor A. Wagner[3,4], Sarah E. Palmer[5], Adam M. Spivak[6], Joshua T. Schiffer[1,2,7*]

[1]Fred Hutchinson Cancer Research Center, Vaccine and Infectious Diseases Division, Seattle, WA, USA
[2]University of Washington, Department of Medicine, Seattle, WA, USA
[3]University of Washington, Department of Pediatrics, Seattle, WA, USA
[4]Seattle Children's Research Institute, Seattle, WA
[5]University of Sydney, The Westmead Institute for Medical Research, Sydney, Australia
[6]University of Utah School of Medicine, Department of Medicine, Salt Lake City, UT, USA
[7]Fred Hutchinson Cancer Research Center, Clinical Research Division, Seattle, WA, USA

*Corresponding author
Email: jschiffe@fhcrc.org (JTS)

Abstract (~150 words)

Antiretroviral therapy (ART) suppresses viral replication in people living with HIV. Yet, infected cells persist for decades on ART. When ART is stopped, HIV replication resumes allowing progression to AIDS. HIV might persist by replicating in drug sanctuaries where ART concentrations are insufficient. Alternatively, latently infected T cells may proliferate. To infer the dominant mechanism, we characterized the relative abundance of unique sequences in ART suppressed individuals. We show that >99% of infected cells are clonal after 1 year of ART, implying cellular proliferation as the primary generator of infected cells. Using a mathematical model, we demonstrate that even if the largest plausible drug sanctuary is assumed, within 6 months of ART nearly all infected cells arise from proliferation rather than HIV replication. Targeting cellular proliferation could reduce the size of the HIV reservoir.

Introduction

Antiretroviral therapy (ART) limits HIV replication in previously uninfected cells leading to elimination of most infected CD4+ T cells.[1] Yet, some infected cells remain and decay at an extremely slow rate despite decades of treatment.[2,3] There is debate whether infection persists due to HIV replication and spread within a small population of cells[4,5] or due to persistence of memory CD4+ T cells with HIV integrated into human chromosomal DNA.[3,6,7] If the latter mechanism predominates, then prolonged cellular lifespan, frequent cellular proliferation, or both sustain stable numbers of infected cells.

To optimize HIV cure strategies, mechanisms sustaining infection must be understood. Persistent viral replication implies a "sanctuary" where ART levels are inadequate and thus the need to improve ART delivery.[8] If HIV persists in a non-replicative state within a *latent reservoir* of memory CD4+ T cells, then the survival mechanisms of these cells are ideal therapeutic targets. Infected cell longevity might be addressed by reactivating the lytic HIV replication cycle,[9] leading to premature cellular demise. Anti-proliferative therapies may limit homeostatic or antigen driven proliferation.[10-12]

The validity of these competing hypotheses hinges on studies of viral evolution. Due to the high mutation rate of HIV reverse transcriptase and the large viral population size,[13] HIV replication in the absence of ART is inferred when the molecular evolution rate exceeds zero.[14,15] New strains accrue synonymous mutations due to the error prone viral replication enzyme, and non-synonymous mutations in response to cytolytic or antibody-mediated immunologic pressure.[16] A recent study documented HIV evolution during months 0-6 of ART in three patients at a rate equivalent to pre-ART time points. New mutations were noted across multiple anatomic compartments, implying widespread circulation of evolving strains,[4] perhaps due to anatomic drug sanctuaries. Alternative explanations are that phylodynamic trees generated during early ART may not consist entirely of sequences within the latent reservoir or were undersampled.[17,18]

In studies of participants on ART for at least one year, viral evolution was not observed despite sampling multiple anatomic compartments.[19-22] Sampled HIV DNA was often linked to historical time points on inferred phylogenetic trees,[23-25] suggesting long-lived latently infected cells as a possible mechanism of HIV persistence.[3,6,7,21,22] Clonal expansions of identical HIV DNA sequences were also present, demonstrating that cellular proliferation generates new infected cells.[4,12,21,26-28] Multiple, equivalent sequences were noted in blood, gut-associated lymphoid tissue (GALT), and lymph nodes, even during the first month of ART.[21,27,28] The majority of reservoir studies relied on sequencing single genes including *env*, *gag* and *pol*, which may overestimate clonality because mutations in other genome segments could go unobserved.[14,29] However, a more comprehensive, whole-genome sequencing approach confirmed abundant replication competent sequence clones.[30] Other studies found multiple, identical HIV integration sites within human chromosomal DNA.[31-35] Given the miniscule probability that two infection events would result in HIV integration within precisely the same human chromosomal locus, this is definitive evidence that infected cells proliferate.[36,37]

The HIV reservoir may be maintained by proliferation with possible contributions from viral replication. However, the ability to identify the origin of persistent infected cells is limited by the fact that current protocols allow sampling of only a tiny fraction of infected cells. Using tools from ecology and mathematical models, we identify that within a year of ART, nearly all new infected cells arise from proliferation.[23,27] Observed HIV evolution during the first 6 months of ART represents a fossil record of past infections rather than persistent viral replication in a drug sanctuary.

Results

Genetic markers of HIV replication versus cellular proliferation. During untreated infection, HIV integrates a single DNA copy into human chromosomal DNA in each infected CD4+ T cell.[38] A majority of new infections are marked by novel mutations due to the high error rate of HIV reverse transcriptase and integration into a unique chromosomal location **(Fig 1)**. Accrual of new mutations during ART would suggest that low levels of replication persist due to inadequate drug delivery to certain micro-anatomic regions.

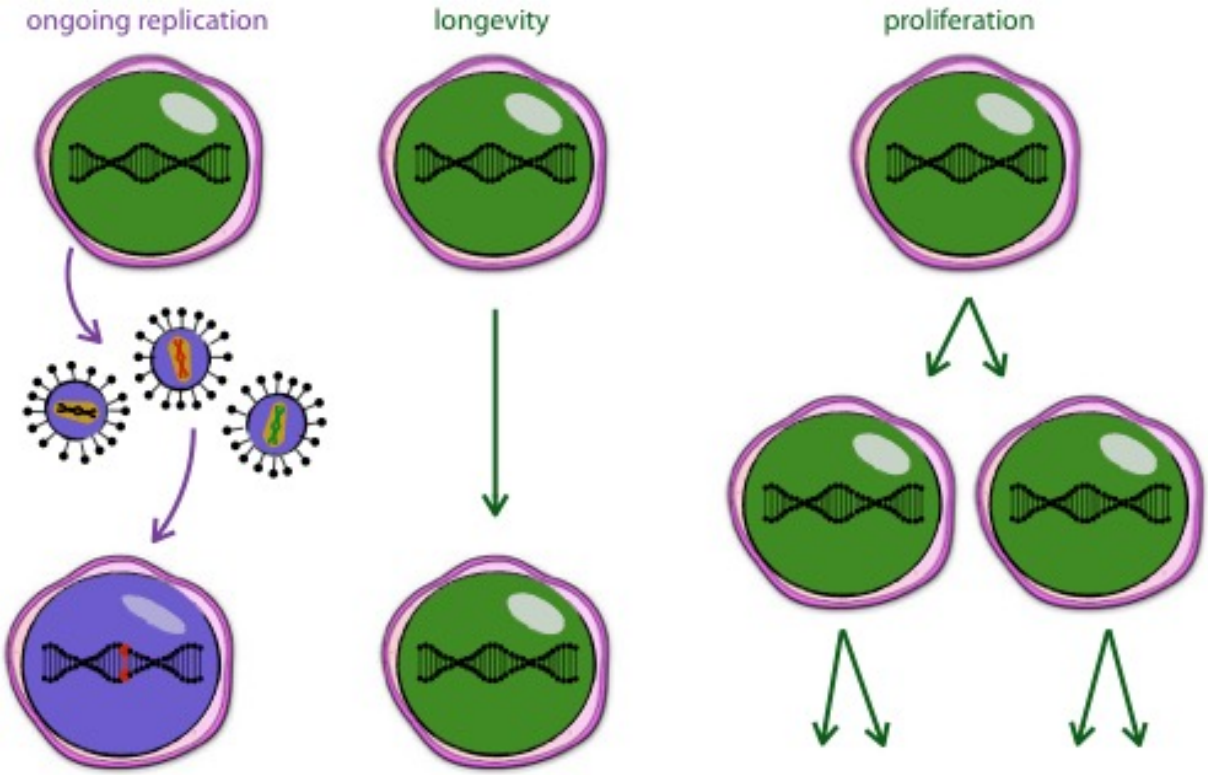

Figure 1. Possible mechanisms for HIV reservoir persistence. *Ongoing viral replication despite ART would lead to accrual of new mutations and novel chromosomal integration sites in newly infected cells, resulting in viral evolution, denoted in this figure by a change in color of the newly infected cell. Alternatively, prolonged persistence of latently infected cells would lead to no new mutations or chromosomal integration sites. Finally, cellular proliferation of latently infected CD4+ T cells would result in multiple cells with equivalent HIV sequences and chromosomal integration sites.*

In a subset of infected CD4+ T-cells, viral replication does not progress beyond chromosomal integration and a single HIV DNA molecule achieves latency.[38] If equivalent HIV sequences with identical chromosomal integration sites are identified in multiple cells, then these viruses were generated via the less error-prone human DNA polymerase during cellular proliferation, rather than HIV replication **(Fig 1)**. We contrast divergent HIV sequences in distinct chromosomal

integration sites (***distinct sequences***) from equivalent sequences in the same chromosomal integration sites (***equivalent sequences***).

Most integrated HIV DNA has accrued mutations that render the virus replication incompetent. Quantification of total HIV DNA copies therefore overestimates the size of the replication competent reservoir by 2-3 orders of magnitude relative to viral outgrowth assays.[39] Replication incompetent, equivalent HIV sequences are commonly present in multiple cells[21,27] and could only have arisen from cellular proliferation. Yet, equivalent replication competent sequences are also found and at comparable frequencies.[30] While total HIV DNA may not predict quantity of replication competent viruses, estimates of clonal frequency using total HIV DNA might be extrapolated to the smaller replication competent reservoir (~10^6-10^8 CD4+ T cells).[40] While ~1-10 replication competent viruses are present in 10^6-10^8 CD4+ T cells, ~100-1,000 HIV DNA molecules can be identified in a similar number of cells.[40] Therefore, total HIV DNA allows a greater sample size for analysis.

Clonal HIV DNA sequences and replication competent sequences are detectable at various time points during ART. Cellular proliferation generates some proportion of infected cells during ART. Wagner *et al.* sampled HIV DNA in three participants at three time points 1.1-12.3 years following ART initiation.[32] Maldarelli *et al.* sampled HIV DNA from five participants at one to three time points 0.2-14.5 years following ART initiation.[31] In these studies, 4-42% (mean: 28%) of sequences were members of ***observed sequence clones*** (**Fig 2A**),[31,32] meaning that HIV DNA was identified in the same chromosomal integration site in at least two cells. The remaining viruses were identified in a specific chromosomal integration site in only one cell (***observed singletons***). The percentage of observed singletons decreased over time in 5 of 6 study participants with longitudinal samples (**Fig 2A**).[32]

Hosmane *et al.* sequenced replication competent HIV isolates from 12 study participants on ART: 0-28% (mean: 11%) of sequences were members of ***observed sequence clones*** (**Fig 2B**).[30] There was a higher number of experimentally detected sequences (N^{obs} or ***sequence sample size***) for total HIV DNA **(Fig 2C)** than for replication competent HIV **(Fig 2D)**. For total HIV DNA, the number of unique observed sequences (R^{obs} or the ***observed sequence richness***) was always less than N^{obs} **(Fig 2C)** due to the presence of clones. Where $N^{obs} > 20$ for replication competent viruses, R^{obs} was always less than N^{obs}, again due to the presence of clones. The lack of detected clones in 3 participants in Hosmane *et al.* may reflect their low sequence sample size. There was a higher R^{obs} as the sequence sample size increased **(Fig 2C, D)**, suggesting that detection of unique clones increases with deeper sampling. Participants with $N^{obs} < 20$ were excluded from individual analyses described below but were included for population level evaluations.

To examine the structure of clonal total and replication competent HIV DNA, we ranked each unique observed sequence according to its proportional abundance. Mathematically, we denote these rank-abundance curves with $a(r)$ such that $a(1)$ is the abundance of the largest clone. Moreover, we calculated the number of observed clones having exactly i identical sequences as $N_i = \sum_r I_i(r)$ where $I_i(r)$ is the indicator function equal to 1 if $a(r) = i$ and zero otherwise. The number of clones having only one member, or "singletons", then is the number of r with $a(r) =$

1, denoted N_1. The fraction of *cells* that were observed in a clone of exactly i sequences is denoted $f_i = iN_i/N$.

For total HIV DNA, at each participant time point, certain sequences predominated: the largest observed sequence clone $a(r = 1)$ contained 3-62 sequences (mean: 13), accounting for 3-32% (mean= 12%) of total observed sequences. The number of observed non-singleton sequence clones $N_{i>1}$ in the 17 samples ranged from 1-83 (mean: 12).

For replication competent HIV DNA in the 5 persons having sample-size $N^{obs} > 20$, certain sequences also dominated: the largest observed clone contained 3-9 sequences (mean: 6.8), accounting for 11-42% (mean= 28%) of total observed sequences. The number of non-singleton sequence clones $N_{i>1}$ in the 5 samples ranged from 1-7 (mean: 3.8).

In both data sets, the morphology of observed rank abundance $a(r)$ was similar across samples **(Fig 2E, F)**, suggesting a generic structure to clonal distribution of reservoir cells.

The total number of HIV DNA sequences in the HIV reservoir exceeds the total number of unique sequences. We next determined that observed singletons are typically members of ***true sequence clones***. Indeed, only sampling every infected cell in the body would provide a proportion of viral sequences that are ***true singletons***.

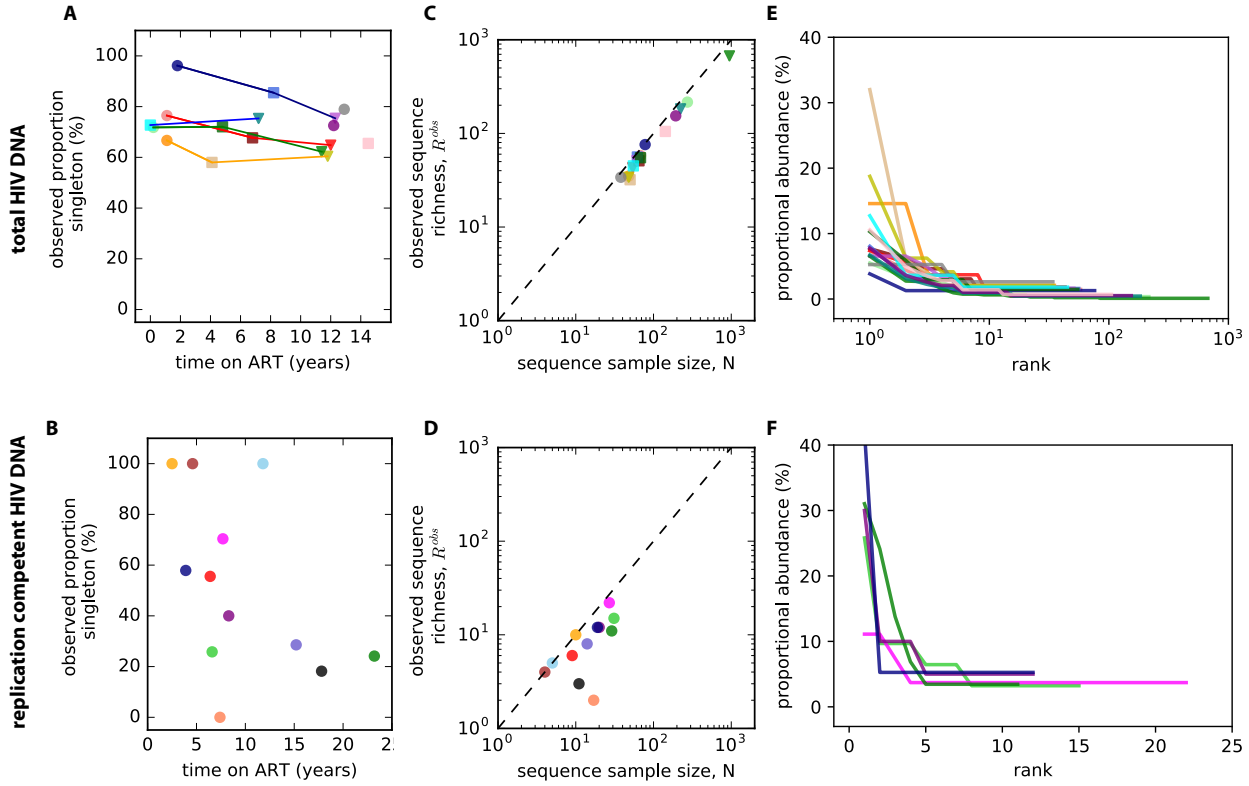

Figure 2. Evidence for clonal HIV sequences. *HIV DNA integration site data was collated from two studies in* ***A, C & E*** *(Wagner et al., and Maldarelli et al.)*[31,32]*. Replication competent HIV DNA whole genome sequence data was gathered from Hosmane et al. in* ***B, D & F.***[30] ***A.*** *The proportion of observed singleton HIV DNA sequences generally decreases with time on ART: the proportion of observed singleton sequences f_1 is listed with color coding for 8 study participants. The remaining proportion of sequences are non-singleton $a(r) > 1$ and likely arose from cellular proliferation. Longitudinal sequences from the same participant at sequential time points on ART are connected with lines.* ***B.*** *The observed proportion of singleton replication competent HIV DNA sequences in 12 study participants is highly variable.* ***C & D.*** *Sample size of HIV DNA* ***(C)*** *and replication competent HIV DNA* ***(D)*** *increases the number of observed unique sequences (observed sequence richness). The number of total sequences at each time point is plotted against the observed sequence richness. For all HIV DNA samples and when $N > 20$ for replication competent HIV DNA, the observed richness is always less than the sample size (to the right of the dotted line y=x), owing to the presence of sequence clones.* ***E & F.*** *Rank abundance curves of HIV DNA* ***(E)*** *and replication competent HIV DNA* ***(F)*** *of sequence clones (only the 5 participants with $N > 20$ are included).*

To address the limited sample size of infected cells during ART, we generated rarefaction curves (**Fig 3A** for total HIV DNA and **Fig 3B** for replication competent HIV when $N^{obs} > 20$) to capture the relationship between sample size and richness: these curves demonstrate a monotonic, nonlinear increase in the number of observed distinct sequences with increased sampling. Repeated sampling would continue to detect new distinct sequences **(Fig 2C, D)**. However, with much greater sampling depth, the rarefaction curves may eventually plateau at the ***true sequence richness***, the number of unique sequences in the entire reservoir. Current sampling depth remains on the steep, initial portion of the curve.

To estimate an absolute lower bound for true sequence richness, we used the Chao estimator, a nonparametric ecologic tool that combines number of observed singletons N_1 and doubletons N_2 (**Fig 2E, F**), with R^{obs} (**Fig 2C, D**).[41] Theoretical values for true richness range from 1 (if all sequences are identical and originate from a single proliferative cell) to N (if all sequences are distinct and originate from error-prone viral replication). The estimated lower bound for true sequence richness exceeded observed richness, usually by an order of magnitude and varied between persons and over time (**Fig 3C** for total HIV DNA and **Fig 3D** for replication competent HIV). These initial lower bound estimates for observed sequence richness were far lower than the total number of HIV DNA and replication competent HIV DNA sequences, suggesting that clones may predominate.

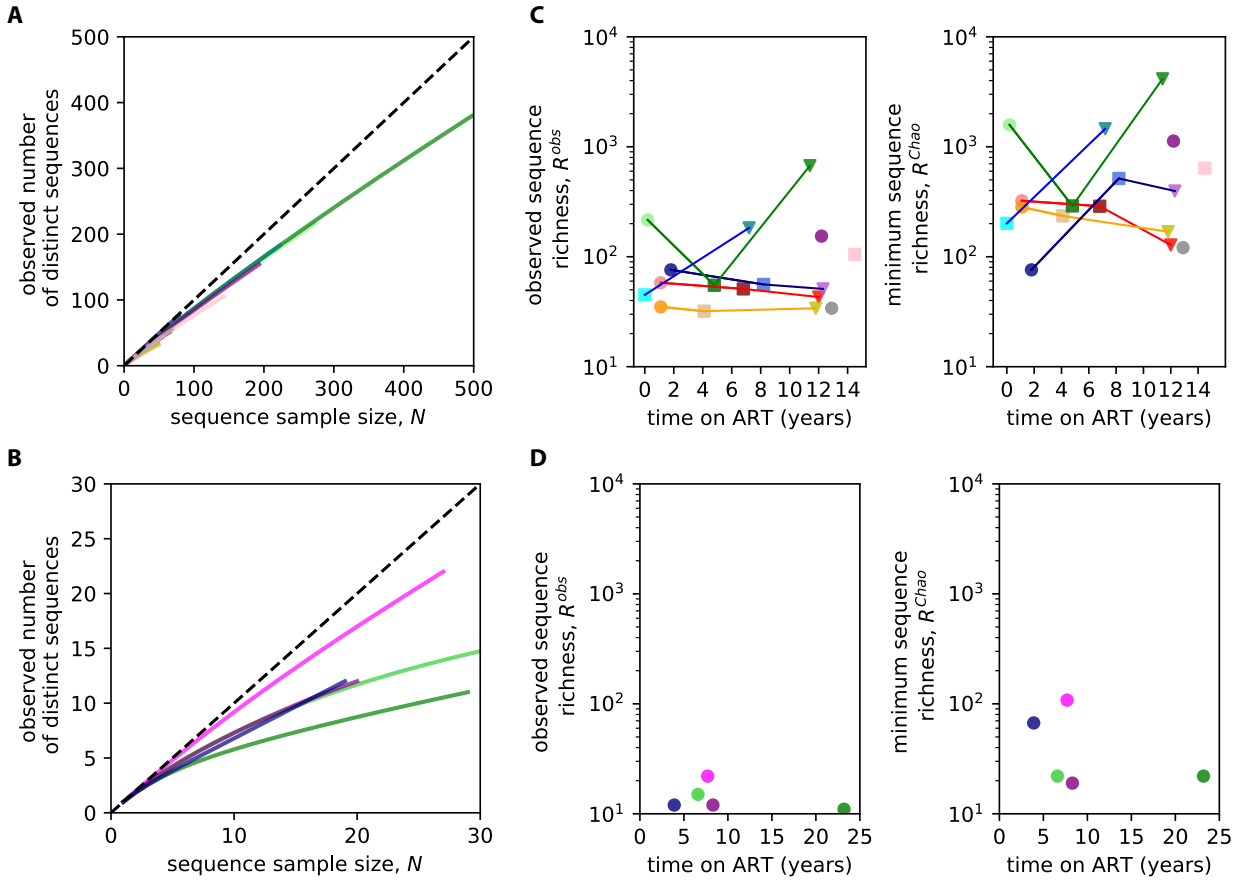

Figure 3. The actual total number of distinct HIV sequences far exceeds the observed total number of distinct HIV sequences during ART. A. *Sample rarefaction curves for all 17 time points from the 8 study participants in* ***Fig 2E*** *demonstrate the observed number of distinct integration sites as a function of HIV DNA sequence experimental sample size.* ***B.*** *Sample rarefaction curves for all 5 study participants in* ***Fig 2F*** *demonstrate the observed number of distinct replication competent HIV DNA sequences as a function of sequence sample size. In both cases, at low sample size, distinct sequences are commonly observed with each new sample. As sample size increases, distinct sequences are increasingly less likely to be detected owing to the presence of repeatedly detected sequence clones. As more and more unique sequences are detected, the curve flattens until all unique sequences are detected and the curve is completely flat.* ***C & D.*** *The observed HIV DNA sequence richness* ***(C)*** *and observed replication competent HIV DNA sequence richness* ***(D)*** *from the sampled data (left panels). The lower bound for true sequence richness estimated using the non-parametric Chao estimator (right panel) is greater than observed richness for each sample. Longitudinal sequences at sequential time points on ART are connected with lines.*

Virtually all observed HIV sequences are members of large proliferative clones. The Chao estimator does not include information about the total population size. However, estimates for the total number of total DNA and replication competent sequences in the entire reservoir exist.[40] Using that additional information, we developed a model to estimate the true rank abundance of the entire HIV reservoir for each participant time point. Based on the observation

that participant data was roughly log-log-linear, we chose a power-law relationship for the theoretical rank abundance curves $a \propto r^{-\alpha}$. Other functional forms were explored (exponential, linear and biphasic power law) but were suboptimal (data not shown).

We fit the model by fixing reservoir size ($L \approx 10^9$ for HIV DNA and $L \approx 10^7$ for intact, replication competent HIV DNA). In **Sup Fig 1** we illustrate our fitting procedure. For each participant time point we generated 2,500 possible models, which varied according to power law exponent α and richness R (**Sup Fig 1A**). We then recapitulated the experiment by taking N^{obs} random samples from each model distribution. Model "fit error" (**Fig 4B**) was calculated as the logarithm of the residual sum of squares between the average of 10 replicate simulations and experimental data in terms of cumulative proportional abundances. The best fit models (top 5 models in **Sup Fig 1B**) minimized this fit error metric**.** Of note, our resampling method accurately inferred the power law exponent unlike log-log-linear regression applied directly to the observed rank abundance curve, which consistently underestimated α (**Sup Fig 1F**).

While α can be identified from the minimum error model, above a certain richness, we cannot identify R (i.e., different richness resulted in equivalently low error **Sup Fig 1C**). Recognizing this uncertainty, we developed an integral approximation to allow for the largest possible richness given L and α (**Sup Fig 1D,** derivation in the **Methods**). Because the Chao estimator (**Fig 3C, D**) provided a lower bound, this approximation fully constrains our estimate of true sequence richness.

Examples of model fit for individual participants are shown in **Fig 4A, B** for total HIV DNA and **Fig 5A, B** for replication competent HIV DNA. Our method demonstrated excellent fit to cumulative proportional abundances of observed clones for total HIV DNA (**Fig 4A**) and replication competent HIV DNA (**Fig 5A**). For total HIV DNA **(Fig 4B)** and replication competent HIV DNA (**Fig 5B**), optimal fit was noted within narrow ranges for the power law slope parameter but across a wide possible range of true sequence richness.

We next generated extrapolated distributions of the entire HIV reservoir for each participant time point with the 5 best fitting parameter sets. We observed similar estimates for the number of the largest clones, which account for approximately 50% of the reservoir (600-900 clones for HIV DNA in **Fig 4C** and 2-7 clones for replication competent HIV DNA in **Fig 5C**). However, the tail of the reservoir, which consists of thousands of smaller clones, varied considerably across the parameter sets with 8,000-200,000 possible clones accounting for 90% of the HIV DNA and 100-2,000 possible clones accounting for 90% of replication competent HIV. This variability reflects the fact that true sequence richness is only partially identifiable using our procedure.

Even under the most conservative assumptions (maximum possible true sequence richness), the vast majority of sequences were predicted to be members of true sequence clones. For the participant in **Fig 4C**, a maximum of 600,000 clones were needed to reach 100% cumulative abundance for HIV DNA. The ratio of estimated true sequence richness to the total number of infected cells R/L with HIV DNA (~10^6: 10^9) represents an upper limit on the fraction of sequences that are true singletons: we estimate that at least 99.9% of infected cells contain

true clonal sequences. Similarly, the ratio of estimated true sequence richness to the total number of infected cells with replication competent HIV for the participant in **Fig 5C** was 10^5: 10^7. Hence, at least 99% of cells contain true clonal sequences. For both HIV DNA (**Fig 4D**) and replication competent virus (**Fig 5D**)**,** the top 1,000 clones were estimated to be large (>10,000 and >100 cells respectively), suggesting that nearly all observed singleton sequences in these participants are likely to be members of large sequence clones. These results argue for proliferation as the major generative mechanism of persistent HIV-infected cells.

We applied our model to generate extrapolated distributions for all study participants in **Fig 2**. We biased against a clonally dominated reservoir to the greatest extent possible by selecting the best fitting power law exponent and then calculating the maximum possible sequence richness. The power law slope parameter was more variable and on average lower across participants for HIV DNA ($\alpha = 1.05 \pm 0.2$) than for replication competent HIV DNA ($\alpha = 1.32 \pm 0.2$). As a result, the predicted cumulative distribution of HIV DNA (**Fig 4E**) was often linear with log rank as compared to concave-down with log rank noted for replication competent HIV DNA (**Fig 5E**), suggesting that a smaller number of extremely large clones might make up a higher proportion of the replication competent HIV reservoir. In all participants, for HIV DNA (**Fig 4F**) and replication competent virus (**Fig 5F**), the top 1,000 clones were large (>10,000 and >100 cells respectively), suggesting that observed singleton sequences are members of large true sequence clones.

To increase sample size and eliminate bias related to excluding participants with low sample sizes, we combined results from all participant time points for HIV DNA (17 time points) and replication competent HIV (12 time points) into single rank order distribution curves. We then fit the power law models to both sets of data (**Sup Fig 2A, 3A**). We again noted a narrow range of possible values for the power law exponent and a large range of possible values for true sequence richness. The exponent was $\alpha < 1$ for total HIV DNA and $\alpha \approx 1$ for replication competent virus (**Sup Fig 2B, 3B**), leading to concave-up and linear relationships between cumulative proportional abundance and log rank, respectively (**Sup Fig 2C, 3C**). We estimated that at least 99.9% of cells with HIV DNA (**Sup Fig 2C**) and 99.8% of cells with replication competent HIV (**Sup Fig 3C**) contain true clonal sequences. The top 1,000 HIV DNA clones (**Sup Fig 2D**) and replication competent clones (**Sup Fig 3D**) contained >10^5 and >10^3 cells respectively.

Using the population level data, we generated sample rarefaction curves from the extrapolated rank abundance curves. These curves demonstrate that even if more than 10,000 HIV DNA or replication competent sequences were obtained from a sample, new unique sequences would continue to be detected (**Sup Fig 4**). Statistical extrapolation of observed data is therefore necessary to estimate the clonal ecology of the HIV reservoir. Even if sample size could be increased 100-fold, samples would be dominated by clones from the "tip of the iceberg."

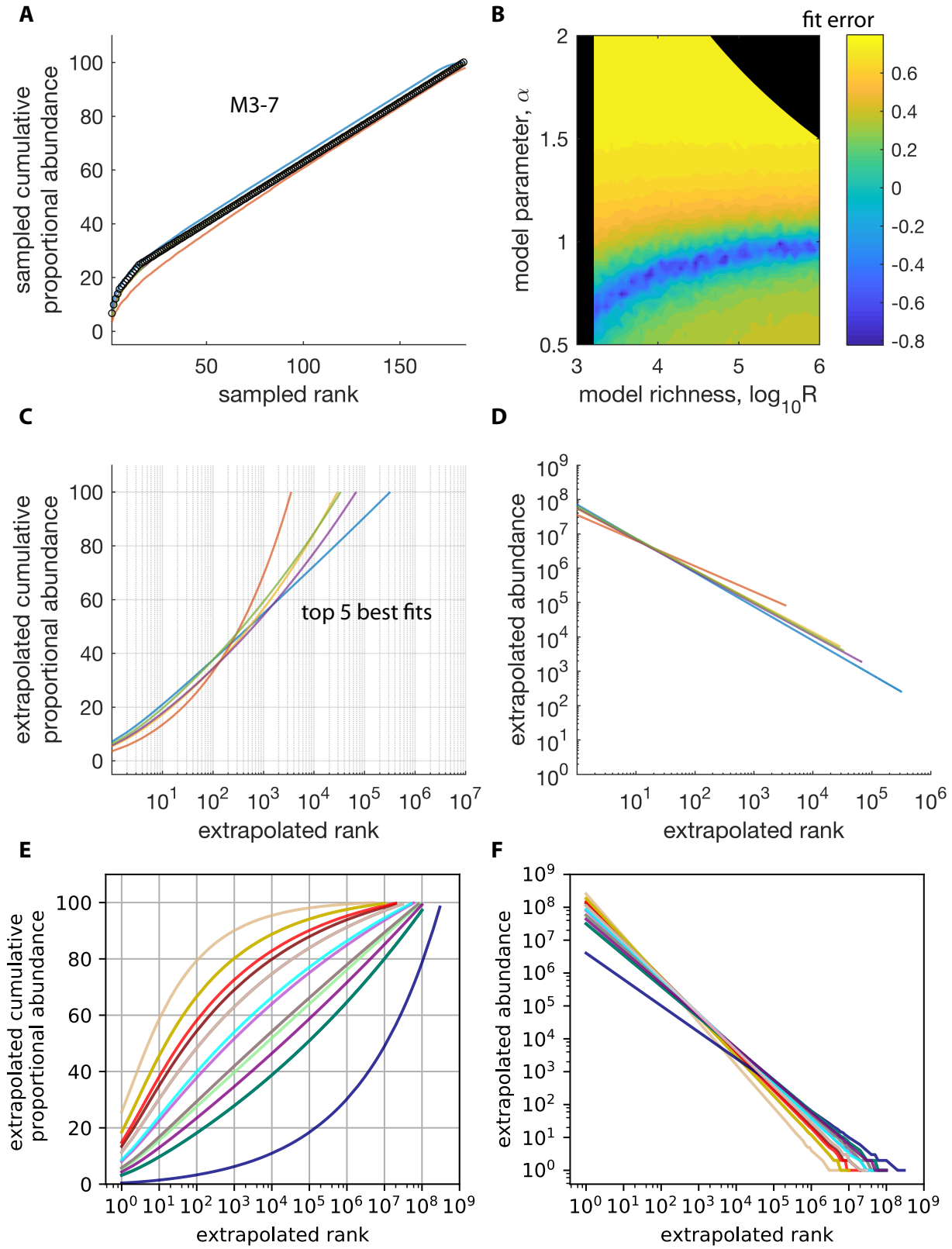

Figure 4. The majority of HIV DNA sequences are members of large sequence clones. *To model the true rank abundance distribution of the HIV reservoir, we use a power law model and*

recapitulated experimental sampling in silico to fit the best model to participant data in ***Fig 2A, C, E****. Theoretical distributions varied according to the slope of the power law and the true sequence richness, but were fixed at 10^9 total HIV DNA sequences.* ***A.*** *Five best model fits to cumulative proportional abundance curves from a single representative participant (M3, 7 years on ART). The actual data is represented by single black circles for each rank; the 5 colored model lines are superimposed based on virtually equivalent fit to the data.* ***B.*** *Heat diagram representing model fit according to power law exponent α and true sequence richness R with best fit noted by minimum error score (blue color, see details of calculation in results above); black shaded areas represent parameter sets excluded based on the Chao estimator (lower bound on sequence richness) and mathematical constraints of the power law (upper bound for sequence richness). A wide range of values for sequence richness allow excellent model fit while the power law exponent is well defined.* ***C.*** *Extrapolations of the best fit cumulative distribution function to the entire pool of 10^9 infected cells; under the most conservative estimates, the top 10^6 ranked clones constitute the entire reservoir.* ***D.*** *Extrapolations of the best fit power law to the entire pool of 10^9 infected cells; the top 1000 clones consist of $>10^4$ cells each.* ***E.*** *Extrapolations of the best fit cumulative distribution function to the entire pool of 10^9 infected cells for all participant time points in* ***Fig 2A****; we assume the maximum possible sequence richness in each case and still note a predominance of sequence clones.* ***F.*** *Extrapolations of the best fit power law to the entire pool of 10^9 infected cell for all participants in* ***Fig 2A;*** *the top 1000 clones each consist of $>10^4$ cells each.*

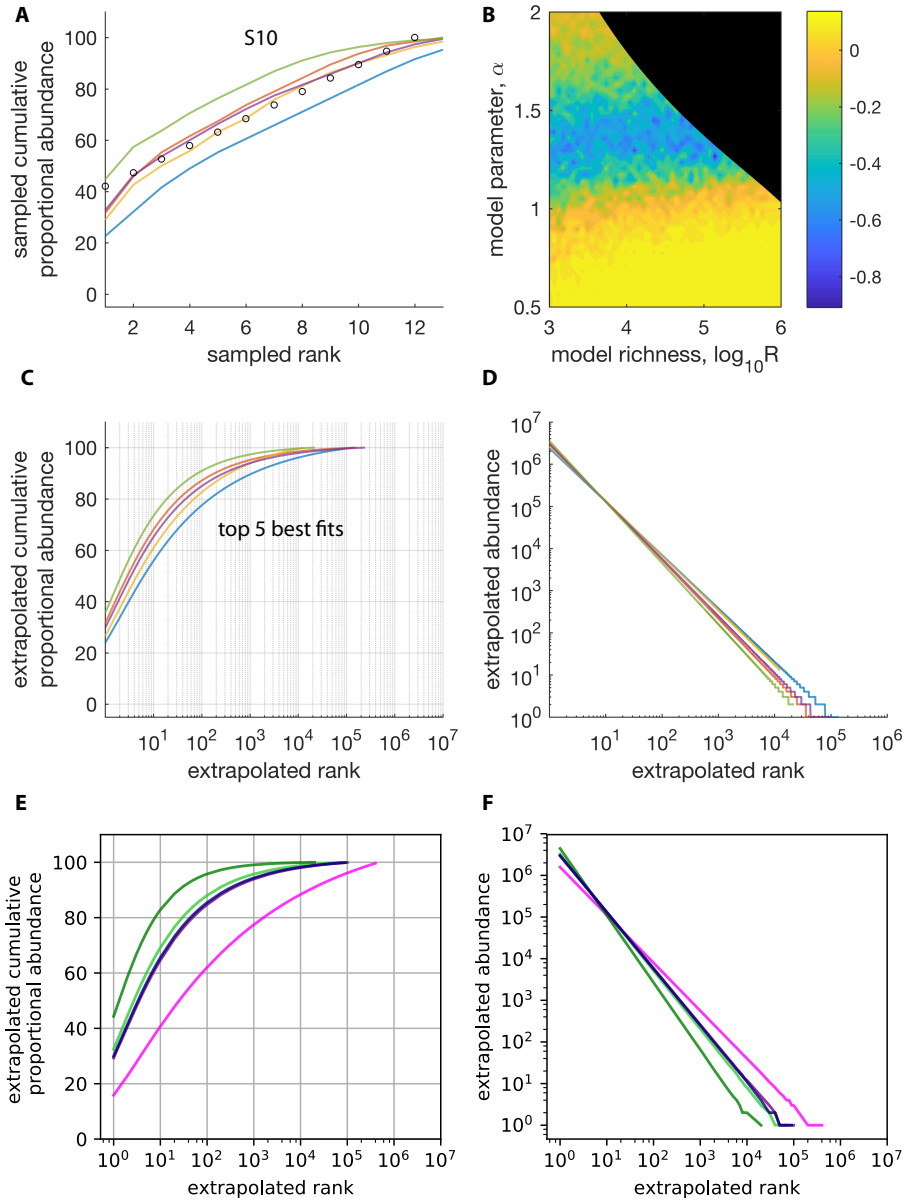

Figure 5. The majority of replication competent HIV sequences in the reservoir are members of large sequence clones. *To recapitulate experimental conditions in* ***Figure 2B, D & F****, we performed in silico sampling (sample size equal to the experimental sample size) from 2,500 theoretical power law distributions of HIV DNA clone size distributions sorted by rank. Theoretical distributions varied according to the exponent of the power law model and the true sequence richness, but were fixed at a reservoir size of* 10^7 *replication competent HIV DNA sequences.* ***A.*** *Five best model fits to cumulative proportional abundance curves from a single representative participant (S10). The actual data is represented by single black circles for each rank; the 5 colored model lines are from five separate parameter sets.* ***B.*** *Heat map representing model fit according to power law slope* α *and true sequence richness* R *with best fit noted by lowest error (blue color); the black shaded area represents parameter sets excluded based on mathematical constraints of the power law (upper bound on sequence richness). A wide range*

of values for sequence richness (<10^5 sequences) allow excellent model fit while power law slope falls within a narrow range. ***C.*** *Extrapolations of the best fit cumulative distribution function to the entire pool of 10^7 infected cells; under the most conservative estimates, the top 10^5 ranked clones constitute the entire reservoir.* ***D.*** *Extrapolations of the best fit power law to the entire pool of 10^7 infected cells; the top 100 clones consist of >10^4 cells each.* ***E.*** *Extrapolations of the best fit cumulative distribution function to the entire pool of 10^7 infected cells for all participants and time points (see original data in* ***Fig 2B****); we assume the largest possible observed sequence richness in each case and still note a predominance of sequence clones.* ***F.*** *Extrapolations of the best fit power law to the entire pool of 10^7 infected cell for all participants in* ***Fig 2B;*** *the top 100 clones again consist of >10^4 cells each.*

Our mathematical model includes both sanctuary and reservoir properties. To reconcile possible evidence of viral evolution during months 0-6 of ART[4] with the high proportions of clonal sequences after one year of ART, we developed a viral dynamic model. Our model (**Fig 6A**) consists of differential equations, described in detail in the **Methods**. The majority of model parameter values are obtained from the literature (**Sup Table 1**).

Briefly, we classify rapid death δ_1 and viral production within actively infected cells I_1. Cells with longer half-life I_2 are activated to I_1 at rate ξ_2. I_2 may represent CD4+ T cells with a prolonged pre-integration phase, but their precise biology does not affect model outcomes.[42] The state $I_{3(j)}$ represents latently infected reservoir cells of phenotype j, which contain a single chromosomally integrated HIV DNA provirus.[38] I_3 reactivates to I_1 at rate ξ_3.[43] The probabilities of a newly infected cell entering $I_1, I_2, I_{3(j)}$, are $\tau_1, \tau_2, \tau_{3(j)}$. Because we are focused on the role of proliferation, we assume sub-populations of I_3,[12] including effector memory (T_{em}), central memory (T_{cm}), and naïve (T_n) CD4+ T cells, which proliferate and die at different rates $\alpha_{3(j)}, \delta_{3(j)}$.[12,36,37] Our model includes the observation that activated, uninfected CD4+ T cells *S*, the critical targets for replicating HIV, decay during ART.[44,45] The decay rate of *S* is slower than concurrent decline in HIV RNA[46-48] because abnormal T cell activation persists for more than a year after ART.[45]

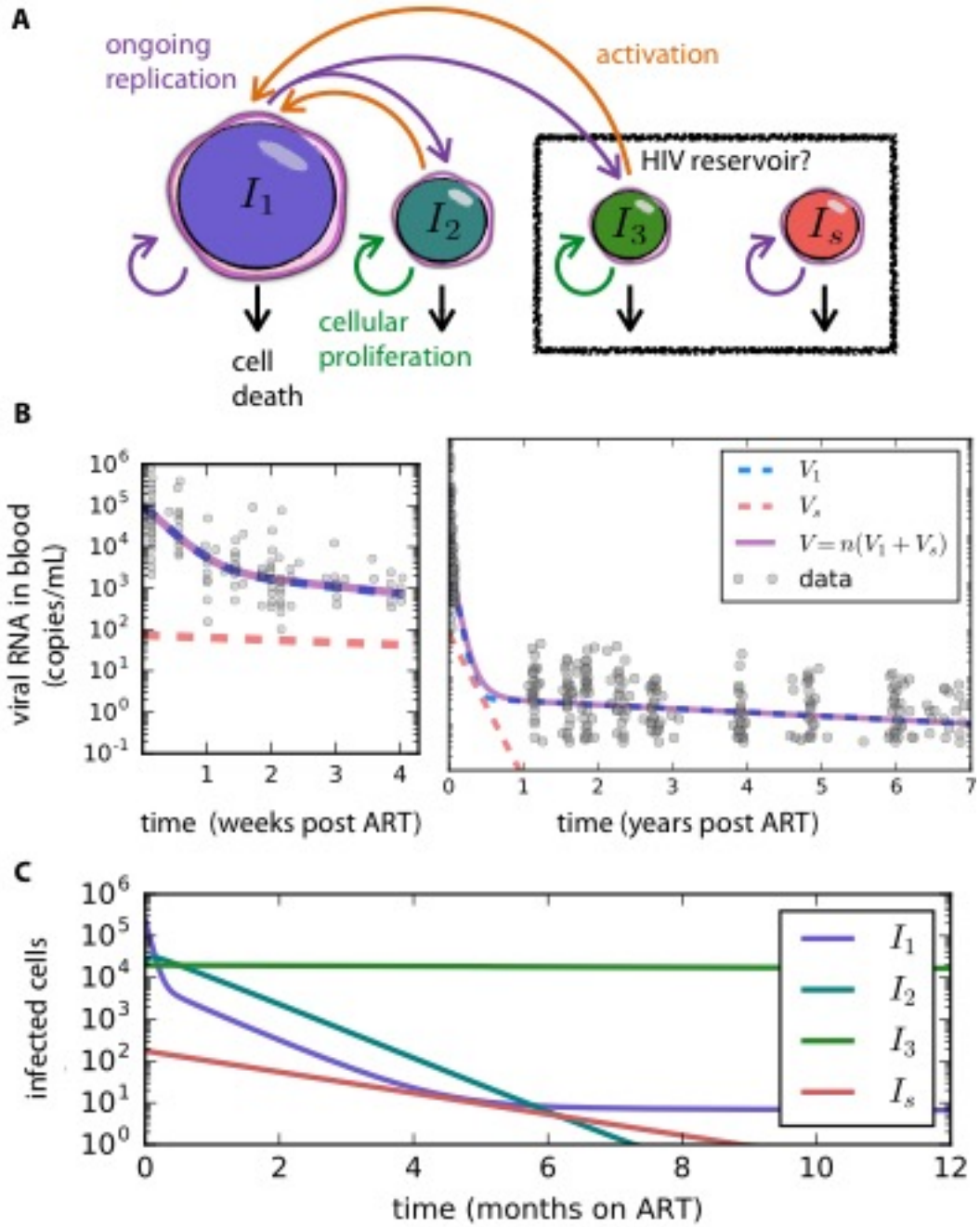

Figure 6. The mathematical model generates three-phase decay. A. *Model structure: newly infected cells become infectious cells with replicating virus I_1, pre-integration latent cells I_2 and post-integration latent cells $I_{3(j)}$, which include a phenotypic index j for effector memory (T_{em}), central memory (T_{cm}) or naïve (T_n) CD4+ T cells that proliferate and die at different rates. Sanctuary cells I_S persistently produce replicating HIV. Parameters and their values are discussed in the* ***Methods*** *and listed in* ***Sup Table 1. B.*** *The mathematical model recapitulates the general three-phase viral decay trend observed in patients. The transition from phase one to two decay is visualized in the left panel during the first month of ART. The transition from phase two to three decay is visualized in the right panel during the first year of ART. Dots are individual patient values from Palmer et al.[47] V_1 is virus derived from I_1 while V_S is derived from I_S.* ***C.*** *I_2*

and I_3 become the predominant cell types early during ART. I_S approximates I_1 between months 4 to 6 if a large drug sanctuary is assumed, but is extinguished within a year.

ART potency $\epsilon \in [0,1]$ characterizes decrease in viral infectivity β due to ART.[46] Other dynamic features of infection such as death rate of infected cells, latent cell proliferation rate and reactivation rates of latent cells, are unchanged on ART. In our simulations, the basic reproductive number becomes $R_0(1-\epsilon)$ on ART and is <1 when $\epsilon > 0.95$, meaning that each cell infects fewer than one other cell and viral load declines from its previous steady state until becoming undetectable. Only short stochastic chains of new infection can occur.

To allow viral evolution despite ART, we included a drug sanctuary (I_S) with an equivalent reproductive number $R_0(1-\epsilon_S)\sim 8$ off of ART. In the drug sanctuary, ART potency is assumed to be negligible ($\epsilon_S = 0$). However, target cell limitation or a local immune response must result in a sanctuary viral set point to prevent infected cells and viral load from growing exponentially which does not occur on ART. The amount of virus produced by the sanctuary V_S is extremely low relative to non-sanctuary regions because ART results in levels undetectable by sensitive assays.[47] We assumed uniform mixing of V_1 and V_S in blood and lymph node homogenates in terms of sampling.[4] Based on observations in treated patients, target cell availability was assumed to decrease on ART at rate ζ,[44,45] decreasing the sanctuary reproductive number as $R_0(1-\epsilon_S)\exp(-\zeta t)$.

The model accurately simulates viral dynamics during ART. Our goal was to recapitulate average three-phase viral clearance dynamics during ART rather than discriminate subtle viral clearance rate differences between treated patients. Therefore, we fit the model to ultra-sensitive viral load measurements collected from multiple participants in Palmer *et al*.[47] We included experimentally derived values for most parameter values (**Sup Table 1**), solving only for ξ_2 and ξ_3 by fitting to viral load. Simulations reproduce three phases of viral clearance (**Fig 6B**) and predict trajectories of infected cell compartments (**Fig 6C**). While a minority of cells infected pre-ART are latently infected (I_3), the transition to I_3 predominance is predicted to occur during month one of ART. If a large sanctuary is assumed with 0.1% of virus being produced in a sanctuary pre-therapy (which would allow persistent detection of low-level, sanctuary-derived viral RNA in blood), then between 4-6 months, I_S approximates I_1 that originate from I_2 and I_3 via reactivation (**Fig 6C**). Sanctuary-derived HIV RNA also approaches reservoir-derived HIV RNA 6 months after ART before contracting to undetectable levels (**Fig 6B**). A small, contracting drug sanctuary provides a possible explanation for why evolving HIV DNA samples are only observed during the first 6 months of ART.[4]

Cellular proliferation sustains HIV infection during ART. We next simulated the model under several plausible sanctuary and reservoir conditions to assess the relative contributions of infection and cellular proliferation in sustaining infected cells. We considered different reservoir compositions based on evidence that effector memory (T_{em}), central memory (T_{cm}) and naïve (T_n) cells proliferate at different rates and that distributions of infection in these cells differ among infected patients.[12,36,37] Because a drug sanctuary has not been observed, its true volume is

unknown and may vary across persons. Thus, we varied the potential size of a sanctuary relative to reservoir cells.

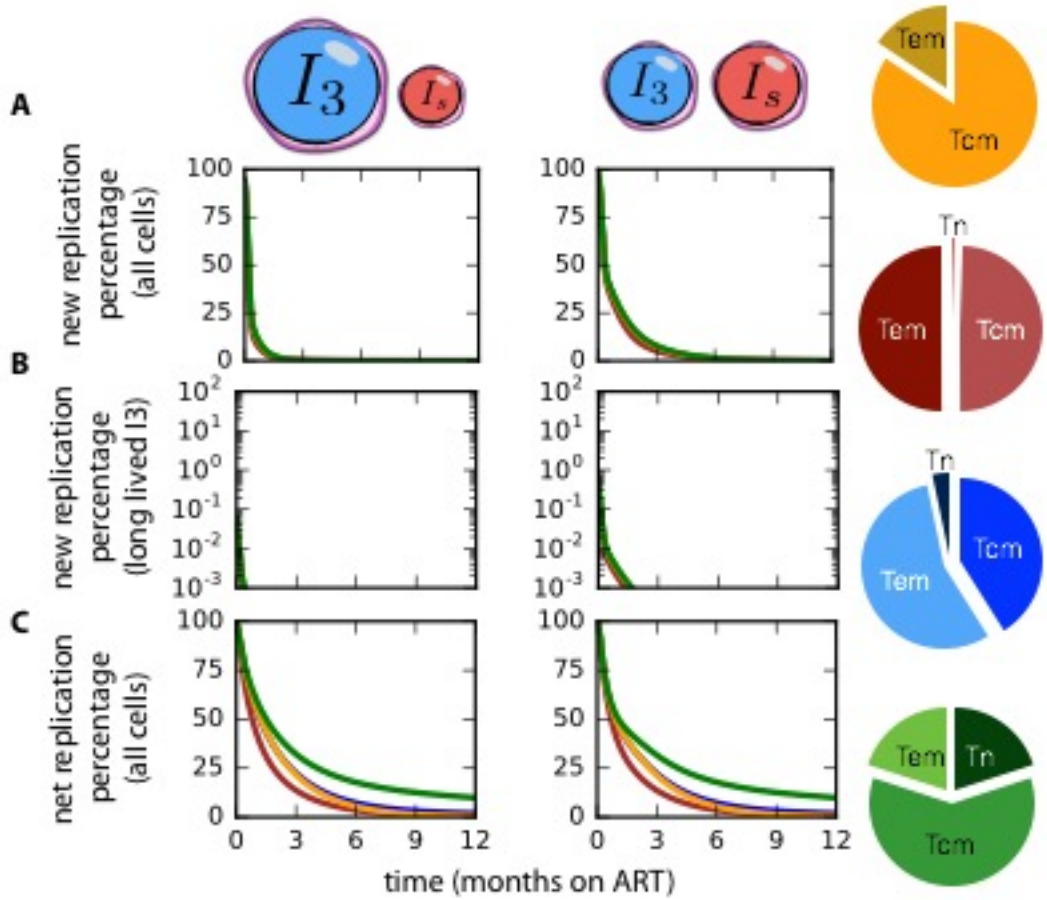

Figure 7. The vast majority of infected cells are generated via proliferation within 6 months of ART initiation. *Different observed proportions of effector memory (T_{em}), central memory (T_{cm}) or naïve (T_n) CD4+ T cells were assumed in the reservoir of simulated patients pre-ART[12] to explore the possible roles of different cellular proliferation rates. Pie chart colors on right correspond with lines in* ***A-C.*** *Different sanctuary and reservoir volumes were assumed. Example distributions of cells with either high or equivalent latent reservoir (I_3) to sanctuary volume (I_S) ratios correspond with panels beneath. Moving left to right, we assume pre-ART reservoir volumes I_3 of 10^7 and 10^6 cells and sanctuary cell frequencies (φ_S) of 1 in 10^6 and 1 in 10^5 of actively infected CD4+ T cells during ART.* ***A.*** *Under all assumptions, once ART is initiated, most new infected cells arise due to cellular proliferation as opposed to HIV replication after 12 months of ART though the transition is slightly more rapid if the sanctuary volume is low relative to the reservoir volume, as in the left-hand column.* ***B.*** *New latently infected reservoir cells (I_3) are generated almost entirely by proliferation soon after ART is initiated under all conditions.* ***C.*** *The observed proportion of infected cells originally generated by HIV infection rather than cellular proliferation will overestimate the actual ongoing proportion during the first 6 months of ART assuming a small or large sanctuary volume.*

Regardless of assumed pre-treatment reservoir composition and sanctuary size, the contribution of sanctuary viruses to generation of new infected cells is negligible after one year of ART. The model predicts that the percentage of newly infected cells generated from viral replication rather than cellular proliferation decreases during ART. This transition is slightly more protracted with a higher ratio of sanctuary to reservoir volume but is not dependent on infected CD4+ T cell phenotypic distribution (**Fig 7A**).

ART immediately eliminates generation of reservoir cells (I_3) by HIV replication (**Fig 7B**). Within days of ART initiation, proliferation is the critical generator of long-lived, HIV-infected cells even if we assume a large drug sanctuary. This finding captures the rarity of infection events on ART even when a sanctuary is assumed, and the frequent turnover of CD4+ T cells. It argues against longevity of infected cells as a driver of reservoir persistence.

Observable HIV DNA sequence evolution during early ART represents a fossil record of prior replication events. Two factors may lead to biased underestimation of infected cell proliferation. First, sample size limitations lead to dramatic overestimation of singletons during ART (**Fig 2**); nearly all observed singletons in persons on suppressive ART for more than a year are likely to be members of populous clones (**Fig 4D, F**).

Second, due to the several-month lifespan of memory CD4+ T cells, observed HIV DNA-containing cells with sequence divergence may reflect historical rather than contemporaneous infection events, particularly during the first 6 months of ART. Under the assumption of a small or large drug sanctuary, the percentage of observed cells generated from viral replication rather than cellular proliferation at 6 months of ART (5-25% in **Fig 7C**) is higher than the percentage of cells currently being generated from HIV replication, which approaches zero (**Fig 7A**). A higher fraction of slowly proliferating T_n cells exacerbates the difference between historical and contemporaneous generation of infected cells (**Fig 7C**). During the first several months of ART, sequences on the most distal branches of a phylogenetic tree could be misinterpreted as *new* infection, when in fact these sequences were generated by an infection event several months beforehand (**Fig 8**). After 12 months of ART, the historical and actual percentage of cells generated from HIV replication converge near zero for all parameter sets—in keeping with our ecological estimation (**Fig 4**). The lag between historical and actual ongoing replication generation is an emergent feature of the model, whether or not a drug sanctuary is included.

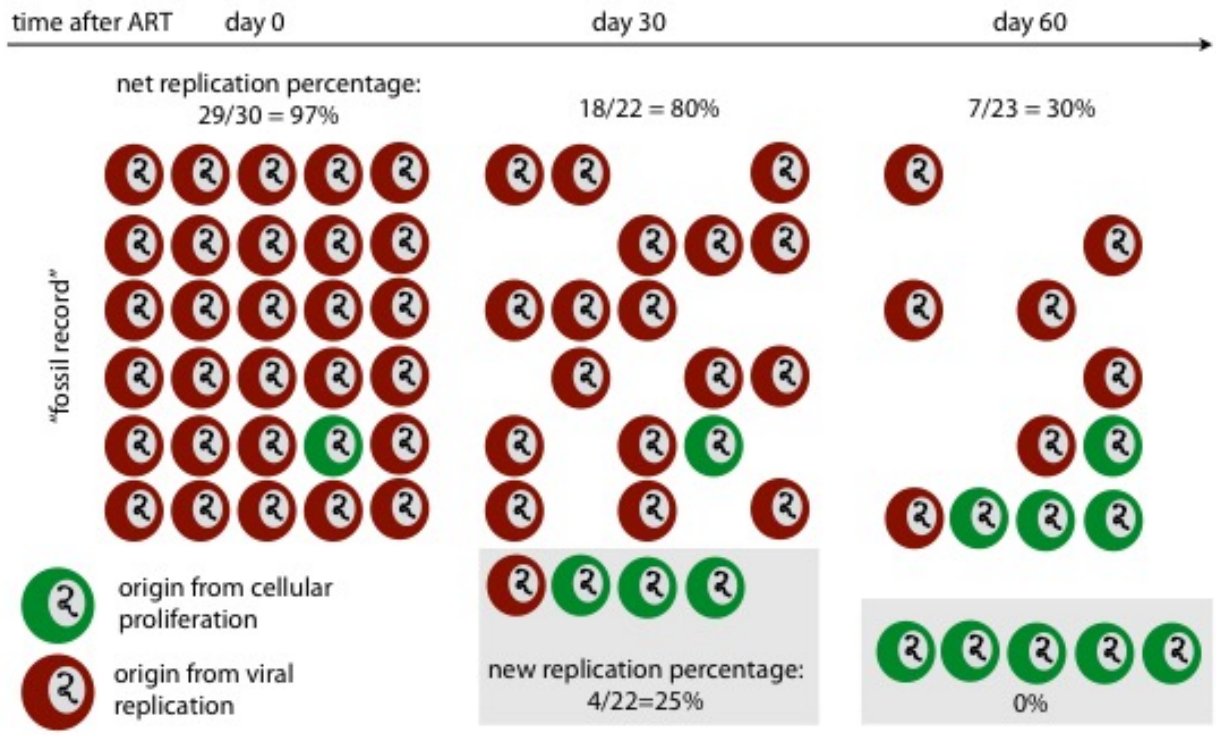

Figure 8. Qualitative illustration of the fossil record phenomenon. *While the proportion of infected cells that were once generated by HIV replication (the net replication percentage, or "fossil record" of HIV replication) remains nonzero for the first few months of ART, the proportion of cells generated by HIV replication on any day during suppressive ART is small (shaded box), and approaches zero even after 2 months of ART. This "fossil record" effect will be evident with or without a drug sanctuary and becomes negligible after a year of ART as in* ***Fig 6.***

Different factors drive observed and ongoing reservoir dynamics during early ART. We next performed sensitivity analyses to identify parameters that predict a more protracted conversion from viral replication to proliferation as a source of new infected cells. Under a wide range of plausible parameter assumptions, new infected cells arise exclusively via proliferation after a year of chronic ART (**Fig 9A**). Only clearance rate of activated CD4+ T cell target cells ζ in the drug sanctuary during ART had an important impact on generation of new infected cells: an extremely slow clearance rate of activated target cells from the drug sanctuary (0.002/day) allowed >1% of new infected cells to be generated from HIV replication rather than proliferation after a year of ART (**Fig 9A**). If this clearance rate was slower than the latent cell clearance rate ($|\zeta| < |\theta_L|$), then the number of cells in the drug sanctuary would ultimately outnumber the number of latently infected cells, but these values are not clinically plausible. Rapid disappearance of HIV replication as a source of new infection is predicted regardless of initial latent cell volume, drug sanctuary volume, and relative fraction of T_{em}, T_{cm}, and T_n.

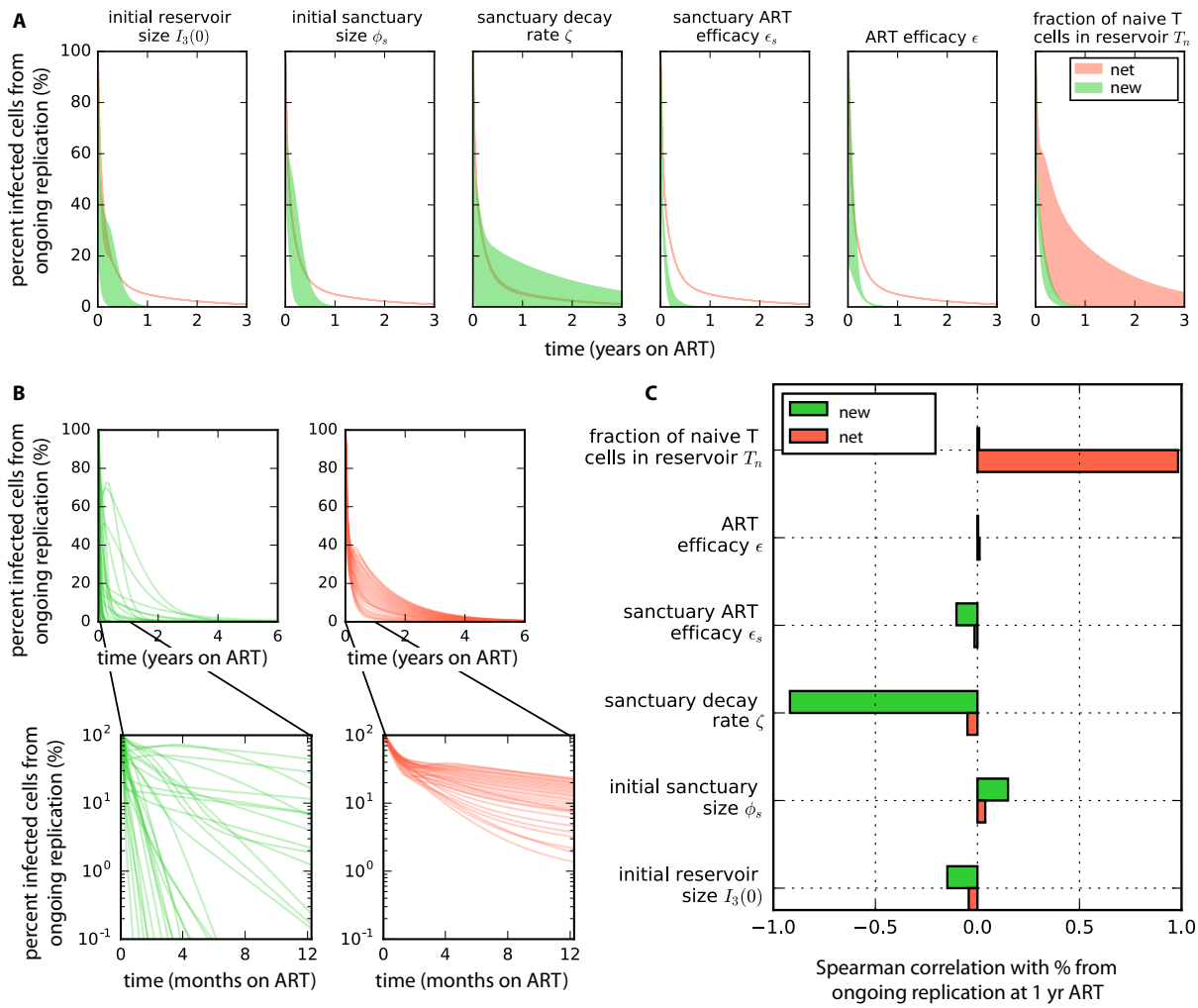

Figure 9. Proliferation of HIV infected cells sustains infection during ART under a wide range of parameter assumptions. **A-C.** See Methods for parameter ranges. ***A.*** *A single parameter local sensitivity analysis (green: new infection events and red: observed infection events) revealed no meaningful difference in percentage of new infected cells generated from viral replication after a year of ART despite variability in initial reservoir volume $I_3(0)$, sanctuary fraction φ_S, and ART effectiveness in and out of the sanctuary (ϵ_S and ϵ). Only an extremely low sanctuary decay rate ζ predicted that a significant percentage (~20%) of infected cells would be generated by HIV replication at one year. Including a high percentage of slowly proliferating naïve CD4+ T cells (T_n) in the reservoir alters the percentage of observed, but not ongoing, percentage of infected cells generated from HIV replication rather than proliferation.* ***B.*** *In a multi-parameter global sensitivity analysis, infection of new cells via HIV replication accounted for fewer than 10% of newly infected cells after a year of ART in most simulations. The plot includes 100 of 1,000 simulated treatments.* ***C.*** *In* ***B,*** *the major driver of a higher percentage of ongoing generation of HIV infected cells via viral replication after one year of ART was low clearance rate of activated CD4+ T cells. The major driver of a higher percentage of observed generation of HIV infected cells via viral replication after*

570 *one year of ART was a higher fraction of naïve T cells in the reservoir. Correlations are measured*
571 *with a Spearman correlation coefficient.*
572
573
574 At time points beyond 6 months of ART, the percentage of observed cells that arose from
575 replication usually exceeded the percentage of infected cells currently arising from replication
576 (**Fig 9A**). The only exception to this trend was when a very low rate of activated CD4+ T cell target
577 cell depletion was assumed (**Fig 9A**). The percentage of observed cells arising from replication
578 rather than proliferation only differed according to percentage of slowly proliferating reservoir
579 cells (T_n) (**Fig 9A**). The drivers of observed and actual ongoing infected cell origin therefore
580 differed completely.
581
582 We simulated 10^4 possible patients with a multi-parameter sensitivity analysis in which all
583 parameter values were randomly selected. These results again showed a rapid transition to
584 proliferation as the source of new infected cells during year one of ART and a higher proportion
585 of observed versus new infection events due to HIV replication (**Fig 9B**). The same parameters
586 correlated with a higher percentage of contemporaneous and observed infection (**Fig 9C**) as in
587 the single parameter sensitivity analysis.
588
589

590 **Discussion**
591
592 To eliminate HIV infected cells during prolonged ART, it is necessary to understand the
593 mechanisms by which they persist. We demonstrate that HIV replication does not maintain
594 infected cell levels after 6 months of ART. Current ART activity is adequate for long-term
595 suppression of replication. Strategies that enhance ART delivery to anatomic drug sanctuaries
596 will not decrease the number of infected cells and are not essential to achieving cure.
597
598 The HIV reservoir also does not persist due to the longevity of latently infected CD4+ T cells. We
599 estimate that >99% of infected cells contain clonal HIV sequences within a year of ART. This trend
600 is recapitulated in our mathematical model, regardless of assumed reservoir composition, or drug
601 sanctuary volume. While many clonal sequences are replication-incompetent, intact, replication-
602 competent viral clones are common.[4,30] Cellular proliferation generates most infected cells,[30] and
603 is a prime therapeutic target.[10,11] Licensed drugs, including mycophenolate mofetil and
604 azathioprine, target lymphocyte proliferation and may deplete the reservoir.[49]
605
606 We demonstrate that observed values of sequence richness and clone size, are substantial
607 underestimates. Current studies only sample the "tip of the iceberg" of the HIV reservoir.
608 Hundreds of thousands of infected cells from a single time point are required to capture true
609 reservoir diversity. This sampling depth could only be achieved as part of an autopsy study.
610
611 The HIV reservoir is defined by a highly structured rank order distribution of relative clone sizes.
612 The best fitting model to this distribution is a power law defined by a narrow range of possible
613 slopes. This distribution specifies a small number of massive clones which comprise a large
614 percentage of observed sequences, but also thousands of much smaller clones. Most observed
615 singleton sequences are likely to be members of clones consisting of millions of cells. The high
616 variance in clone size likely has a mechanistic basis. It is plausible, though unproven, that large
617 clones result from rapid bursts of CD4+ T cell proliferation due to cognate antigen recognition or
618 HIV integration into tumor suppression genes.[31,32] Slower homeostatic proliferation might
619 generate smaller clones.
620
621 A fossil record of infection emerges early during ART. This is an inevitable emergent property of
622 our model, regardless of whether a drug sanctuary is included in the equations. Single sequences
623 that appear to be diverging from the founder virus likely represent a record of prior replication
624 events rather than ongoing infection. Because time of detection does not correlate linearly with
625 sequence age, inference of evolution early during ART is problematic. The fossil record is likely to
626 be a transient effect: within a year of effective ART, observed phylogenetic data likely reflects a
627 lack of contemporaneous HIV evolution.[19-25,27,28,31,32]
628
629 We cannot discriminate whether a small drug sanctuary exists during early ART. The small
630 numbers of true singleton sequences may represent smoldering remnants of previous clonal
631 expansions or ongoing infection. However, a sanctuary must represent a miniscule proportion of
632 infected cells after 6 months of ART to account for lack of persistently detected plasma viremia.
633 Our sensitivity analysis shows that even an extremely low estimate for decrease in activated CD4+

634 T cells will eliminate the sanctuary. In the absence of ART, HIV fuels infection by drawing target
635 cells to sites of viral replication, leading to more rapid cycles of CD4+ T cell turnover.[44] If ART
636 reverses this process, even at a slower rate than empirically demonstrated,[44,45] then the drug
637 sanctuary becomes trivial after several months of ART. Nevertheless, chronic low-grade
638 inflammation does persist on ART, and a small, inconsequential drug sanctuary could endure.
639
640 Phylogenetic data may be misleading in ascribing the origin of infected cells during early ART. The
641 major variable that predicts a higher contribution of *observed* evolution (a larger proportion of
642 slowly proliferating CD4+ T cells in the reservoir) is not the same variable that predicts *actual*
643 proportion of ongoing replication in the reservoir (a slower decrease in availability of CD4+ T cell
644 targets). Cross-sectional observations of reservoir composition may not be as informative as
645 therapies which perturb reservoir size or composition. Assessments of reservoir dynamics are
646 likely to be most reliable if performed after one year of ART.
647
648 We conclude that nearly all infected cells arise from proliferation after the first year of ART. HIV
649 eradication strategies should focus on limiting proliferation.
650
651

Acknowledgments

DBR thanks JM Drake for originally suggesting the Chao estimator. DBR is supported by a Washington Research Postdoctoral Fellowship. JTS is supported by NIH grants: P01 AI030371-24, U19 AI113173-02, R01 AI121129-01, UM1 AI126623, UM1 AI068635. ERD is supported by the National Center for Advancing Translational Sciences of the NIH under Award Number KL2 TR002317.

Methods

Computational code for all calculations and simulations can be found at https://github.com/dbrvs/reservoir_persistence. Sequence data was obtained from the Retrovirus Integration Database (RID).[50]

Rank abundance of HIV integration sites. We used an ecological framework to study latent cells infected with HIV using integration site data. Cellular DNA found with HIV integrated into different integration sites in the human genome were defined as distinct "clones". The number of times a cell was found with the same integration site added to the "abundance" of that clone. By ordering, or "ranking", the clones from largest to smallest by abundance, we developed the rank abundance curve of the experimental data. We expressed this as $r(a)$. A rank abundance curve was created for each study participant time point (time relative to initiation of ART) so that we made no assumptions about the stability or dynamics of the reservoir rank abundance over time.

In our analysis of data from Wagner *et al.*, we combine measurements taken closely in time and use the median time point as done in that published paper. In our analysis of Maldarelli *et al.*, we group by integration sites, or nearest integration site if integration site was not detected. The replication competent sequences of Hosmane *et al.* were counted manually using the phylogenetic trees in that paper.

Calculation of rarefaction curves. We developed rarefaction curves to estimate the expected number of species that would be observed given a certain number of samples from the observed data. To do so, we also defined the clonal abundance by calculating the number of observed clones having exactly i identical sequences as $N_i = \sum_r I_i(r)$ where $I_i(r)$ is the indicator function equal to 1 if $a(r) = i$ and zero otherwise. Then, the expected number of distinct clones that would be observed after k samples given a total sample size of N is

$$R^{obs} - \binom{N}{k}^{-1} \sum_{i=1}^{R^{obs}} \binom{N-N_i}{k}, \qquad (1)$$

where the parentheses indicate combinatorial factors $\binom{x}{y} = x!/y!\,(x-y)!$ often articulated as "x choose y". Later, we extrapolated rarefaction curves using the modeled distributions and a large total reservoir size $L \sim 10^7 \rightarrow 10^9$. Because the number of samples we allowed is orders of magnitude smaller than the number of cells in the reservoir, we can approximate the combinatorial factors as $\binom{x}{y} \sim x^y$ such that now using the true richness R, the expected number of clones observed that would be observed after k samples is

$$R - \sum_{i=1}^{R} 1 - \left(\frac{N_i}{L}\right)^k, \qquad (2)$$

avoiding computation of large factorials.

Nonparametric estimation of species richness. We employed the validated ecological estimator of Chao to estimate a lower bound on the number of distinct HIV sequences (richness) from published data.[51] The "Chao estimator" does not require the true total number of HIV infected cells, only the number of observed singleton N_1 and observed doubleton N_2 clones.[51] It is not a mechanistic model and requires no free parameters. A bias-corrected estimator (avoiding division by zero) for a lower bound on the richness R^{Chao} is expressed

$$R^{Chao} = R^{obs} + \frac{N_1(N_1-1)}{2(N_2+1)}. \quad (3)$$

Confidence intervals can be calculated by defining the ratio of singleton to doubleton sequences as $f_{1:2} = N_1/N_2$, and calculating the variance:[51]

$$\sigma^2 = R^{obs} + N_2\left(\frac{{f_{1:2}}^4}{4} + {f_{1:2}}^3 + \frac{{f_{1:2}}^2}{2}\right), \quad (4)$$

so that 2 standard deviations (~95% confidence intervals) are given by 2σ. These are small in general relative to R^{Chao} so we did not include these values in our plots. Other estimators ("jackknife" 1 and 2) were tested but provided similar and consistently lower estimates of richness than the Chao estimator. Lower estimates of richness increased our confidence that the reservoir is made up of a few large clonal populations rather than many singleton species.

Parametric models to extrapolate the reservoir species abundance curves. Estimates of the size of the HIV reservoir (both replication competent and total) were gathered from the published literature.[40] We then developed a parametric model to quantify the true rank abundance distribution of the complete HIV reservoir. Examination of the data indicated a possible log-log-linear relationship, so we chose a discrete integer power law model so that the probability of a rank is described by $p(r) = \phi(R)r^{-\alpha}$ where the coefficient $\phi(R) = \sum_{r=1}^{R} r^{-\alpha}$ is the normalization constant for the power law. Then, to describe the true rank abundance $a(r)$ we chose the reservoir size depending on the model context (replication competent $L = 10^7$ or total HIV DNA $L = 10^9$). To ensure integer number of cells, we rounded this distribution, and forced the total number of cells to equal the reservoir size. That is,

$$a(r; \alpha, R, L) = |[L\phi(R)r^{-\alpha}]| \quad (5)$$

where $|[]|$ indicates rounding to the nearest integer. Thus, our model depended on two free parameters, a power law exponent α, and the reservoir richness R. Other functional forms were explored but resulted in worse fitting results.

Fitting the model to experimental data. Using the experimental data we found the best-fit model using the following procedure. We fixed the reservoir size L depending on the model context (replication competent or total HIV DNA). We chose a value for R and α from ranges $R \in [1,10^6]$ and $\alpha \in [0,3]$ to specify the model. Then, we sampled the extrapolated distribution 10 times using multinomial sampling with the same number of samples as the experimental data being fit, $\mathcal{M}(N^{obs}, p(r))$. This procedure assumes that sampling cells does

not change the distribution of the reservoir, which is reasonable given the reservoir size. The sampled data were arranged into a single rank abundance curve, averaging over all 10 replicate samplings, and compared to the experimental data in terms of cumulative proportional abundance using residual sum of squares.

2,500 possible parameter sets were made, which resulted in the heat maps in **Figs 4, 5** for example. Because the procedure becomes computationally expensive as $R > 10^7$, we did not explore all possible richness. In theory if all clones had a single member $R = L$, and for the total DNA reservoir, this value would be 10^9. However, we saw that as richness increased, the model was no longer sensitive to R, thus finding the best fit α. We validated the model in **Sup Fig 1** by simulating a reservoir with known power law exponent. This validation shows that simply fitting a power law to the experimental data (using log-log-linear regression) without the extra sampling step necessarily underestimates the power-law exponent, demonstrating the utility of our approach.

We excluded models where $R < R^{Chao}$, but we also sought to identify an upper bound for R. To do so, we used the value of α where R was unidentifiable, usually beyond $R > 10^5$, the value of α was constant. Then, we noted that for large R it is a reasonable approximation to allow $\sum_{r=1}^{R} a(r) = \int_1^R a(r)\, dr$. With α and L specified, R can only be so large, and we solved for the maximal value or the upper bound on the richness given the best fit α and the chosen L. This upper bound then provided the most diverse reservoir, or the reservoir with the most evidence of HIV replication as opposed to proliferation. We then used the maximum richness to ensure we were biasing the results as strongly as possible against our own hypothesis. Finally, we noted that extrapolating abundance curves has been criticized, so we emphasize our attempt to design a simple parametric model was based on the additional information of reservoir size and our goal to define an upper limit on reservoir richness.[52]

Mechanistic model for the persistence of the HIV reservoir. The canonical model for HIV dynamics describes the time-evolution of the concentrations of susceptible $\boldsymbol{S}$ and infected $\boldsymbol{I}$ CD4+ T cells and HIV virus $\boldsymbol{V}$.[46,48,53] Our model grows from the canonical model, simplifying with several approximations and extending the biological detail to simulate HIV dynamics on ART, including a long-lived latent reservoir and a potential drug sanctuary. We describe our assumptions here.

Perelson *et al.* first noticed and quantified a 'biphasic' clearance of HIV virus upon initiation of ART and showed that viral half-lives of 1.5 and 14 days correspond with the half-lives of two infected cell compartments.[46,48] With longer observation times and single-copy viral assays, Palmer *et al.* found four-phases of viral clearance after initiation of ART.[47] Because of uncertainty in distinguishing the third and fourth phase in that study, we focus on the first three decay rates and corresponding cellular compartments, attributing a mixture of the third and fourth phase decay to the clearance of the latent reservoir (half-life 44 months) as measured by Siliciano *et al.* and recently corroborated by Crooks *et al*.[2,3] We developed a mechanistic mathematical model that has three types of infected cells I_1, I_2, I_3 that are meant

786 to simulate productively infected cells, pre-integration infected cells, and latently infected cells,
787 respectively.
788
789 A recent hypothesis about reservoir persistence suggests there may be a small, anatomic
790 sanctuary (1 in 10^5 infected cells) in which ART is not therapeutic.[4] Thus, we included the state
791 variable $\boldsymbol{I_S}$ that is maintained at a constant set-point level prior to ART, where all new infected
792 cells arise from ongoing replication. We opted for this simplification because it biased against
793 our conclusions.
794
795 Next, many studies demonstrated that HIV accelerates immunosenescene through abnormal
796 activation of T cells.[54-56] However, ART results in a marked reduction of T cell activation and
797 apoptosis, a potential signature of HIV susceptible cells.[57] By examining the decline of activation
798 markers for CD4+ T cells, we approximated the decay kinetics of activated T cells upon ART,
799 inferring approximate decay kinetics of the target cells in our model.[44,45,58] A range of initial
800 values exists (from ~5−20% activation) depending on stage of HIV infection, yet after a year of
801 ART, a large percentage of patients return to almost normal, or slightly elevated T cell activation
802 levels (2-3%).[44] Because we assume that target cell depletion is minimal at viral load set-point,
803 we can approximate that the susceptible cell concentration decreases over time as the immune
804 activation decreases, i.e., $S = S(0)e^{-\zeta t}$. This single exponential decay is simplified (it may be
805 biphasic but the data are not granular enough to discriminate this dynamic subtlety). From
806 existing data, the decay constant should be in the range $\zeta \sim [0.002, 0.01]$.[44,58] We extend this
807 decay into the sanctuary, allowing the number of susceptible cells over the whole body to
808 decrease so that we have $I_S = I_1(0)\varphi_S e^{-\zeta t}$ where φ_S is the fraction of infected cells that are in
809 a sanctuary.
810
811 Last, we use the quasi-static approximation that virus is proportional to the number of actively
812 infected cells in all compartments $V = n(I_1 + I_S)$ where $n = \pi/\gamma$, the ratio of the viral
813 production rate to the viral clearance rate (**Sup Table 1**). The model is thus
814

$$
\begin{aligned}
\dot{I}_1 &= \tau_1 \beta_\epsilon S n (I_1 + I_S) - \delta_1 I_1 + \xi_2 I_2 + \textstyle\sum_j \xi_{3(j)} I_{3(j)} \\
\dot{I}_2 &= \tau_2 \beta_\epsilon S n (I_1 + I_S) + (\alpha_2 - \delta_2 - \xi_2) I_2 \\
\dot{I}_{3(j)} &= \tau_{3(j)} \beta_\epsilon S n (I_1 + I_S) + (\alpha_{3(j)} - \delta_{3(j)} - \xi_{3(j)}) I_{3(j)},
\end{aligned}
\tag{6}
$$

815
816
817
818
819 where we use the over-dot to denote the time derivative.
820
821 **Comparing proliferation to ongoing replication: 'net' and 'new' percentages.** Using the ODE
822 model (**Eq 6**), we calculated the total number of newly infected cells generated at time *t* by
823 ongoing replication, both for all infected cells, and model and within the latent reservoir I_3. This
824 quantity we refer to as the "new" percentage due to replication:
825

$$
\frac{\sum_{i,j} \beta_\epsilon S n (I_1 + I_S) \tau_{i(j)}}{\sum_{i,j} \beta_\epsilon S n (I_1 + I_S) \tau_{i(j)} + \alpha_{i(j)} I_{i(j)}} \times 100\%. \tag{7}
$$

826
827

which represents the fraction of newly generated cells due to replication as compared to all newly generated cells. The new replication percentage within the long-lived reservoir is similarly calculated at any time *t*

$$\frac{\sum_j \beta_\epsilon Sn(I_1+I_S)\tau_{3(j)}}{\sum_j \beta_\epsilon Sn(I_1+I_S)\tau_{3(j)} + \alpha_{3(j)}I_{3(j)}} \times 100\%. \quad (8)$$

We also calculated the net percentage that arose at one time due to replication. To biad against our hypothesis, we assumed that all infected cells at the initiation of ART were generated by ongoing replication and then solve for the number of cells whose origin was proliferation:

$$\dot{I}_{i(j)}{}^{pro} = \alpha_{i(j)}I_{i(j)} - (\delta_{i(j)} - \xi_{i(j)})I_{i(j)}{}^{pro}, \quad (9)$$

and similarly solve for those cells whose origin was ongoing viral replication:

$$\dot{I}_{i(j)}{}^{rep} = \tau_{i(j)}\beta_\epsilon Sn(I_1 + I_S) - (\delta_{i(j)} - \xi_{i(j)})I_{i(j)}{}^{rep}. \quad (10)$$

Then the net infection percentage can be written

$$\frac{\sum_{i,j} I_{i(j)}{}^{rep}}{\sum_{i,j} I_{i(j)}{}^{rep} + \sum_{i,j} I_{i(j)}{}^{pro}} \times 100\%, \quad (11)$$

and similarly, the percentage in the long-lived reservoir was calculated separately by removing the summation over the compartment index i, or

$$\frac{\sum_j I_{3(j)}{}^{rep}}{\sum_j I_{3(j)}{}^{rep} + \sum_j I_{3(j)}{}^{pro}} \times 100\%. \quad (12)$$

Sensitivity analysis. Using the parameter ranges $I_3(0) = [0.02,2]$, $\varphi_S = [10^{-6}, 10^{-4}]$, $\zeta = [0.002,0.2]$, $\epsilon = [0.9,0.99]$, $\epsilon_S = [0,0.9]$, $\varrho_j = [0,0.5]$ we completed a local and global sensitivity analysis. For the local analysis, we fixed all values as in **Sup Table 1** and modified one parameter at a time over each listed range. The global analysis was performed by using 10^4 Latin Hypercube samplings of the complete 6-dimensional parameter space.[59] The key outcome, the percent of infections (net and new) due to ongoing replication at 1 year of ART, was correlated to each parameter using the Spearman correlation coefficient—defined by the ratio of the covariance between the outcome and the variable divided by the standard deviations of each when the variables were rank-ordered by value.

References

1. Volberding, P.A. & Deeks, S.G. Antiretroviral therapy and management of HIV infection. *Lancet* **376**, 49-62 (2010).
2. Crooks, A.M.*, et al.* Precise Quantitation of the Latent HIV-1 Reservoir: Implications for Eradication Strategies. *J Infect Dis* **212**, 1361-1365 (2015).
3. Siliciano, J.D.*, et al.* Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med* **9**, 727-728 (2003).
4. Lorenzo-Redondo, R.*, et al.* Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature* **530**, 51-56 (2016).
5. Gunthard, H.F.*, et al.* Evolution of envelope sequences of human immunodeficiency virus type 1 in cellular reservoirs in the setting of potent antiviral therapy. *J Virol* **73**, 9404-9412 (1999).
6. Finzi, D.*, et al.* Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* **278**, 1295-1300 (1997).
7. Finzi, D.*, et al.* Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med* **5**, 512-517 (1999).
8. Fletcher, C.V.*, et al.* Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues. *Proc Natl Acad Sci U S A* **111**, 2307-2312 (2014).
9. Archin, N.M.*, et al.* Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature* **487**, 482-485 (2012).
10. Chapuis, A.G.*, et al.* Effects of mycophenolic acid on human immunodeficiency virus infection in vitro and in vivo. *Nat Med* **6**, 762-768 (2000).
11. Garcia, F.*, et al.* Effect of mycophenolate mofetil on immune response and plasma and lymphatic tissue viral load during and after interruption of highly active antiretroviral therapy for patients with chronic HIV infection: a randomized pilot study. *J Acquir Immune Defic Syndr* **36**, 823-830 (2004).
12. Chomont, N.*, et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med* **15**, 893-900 (2009).
13. Maldarelli, F.*, et al.* HIV populations are large and accumulate high genetic diversity in a nonlinear fashion. *J Virol* **87**, 10313-10323 (2013).
14. Zanini, F.*, et al.* Population genomics of intrapatient HIV-1 evolution. *Elife* **4**(2015).
15. Shankarappa, R.*, et al.* Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* **73**, 10489-10502 (1999).
16. Lemey, P.*, et al.* Synonymous substitution rates predict HIV disease progression as a result of underlying replication dynamics. *PLoS Comput Biol* **3**, e29 (2007).
17. Kearney, M.F.W., A.; Shao, W.; McManus W.R.; Bale, M.J.; Luke, B.; Maldarelli, F.; Mellors, J.M.; Coffin, J.M. Ongoing HIV Replication During ART Reconsidered. *Open Forum Infectious Diseases* **4**, ofx173 (2017).
18. Rosenbloom, D.I.S., Hill, A.L., Laskey, S.B. & Siliciano, R.F. Re-evaluating evolution in the HIV reservoir. *Nature* **551**, E6-E9 (2017).

907 19. Evering, T.H., *et al.* Absence of HIV-1 evolution in the gut-associated lymphoid tissue from
908 patients on combination antiviral therapy initiated during primary infection. *PLoS Pathog*
909 **8**, e1002506 (2012).
910 20. Frenkel, L.M., *et al.* Multiple viral genetic analyses detect low-level human
911 immunodeficiency virus type 1 replication during effective highly active antiretroviral
912 therapy. *J Virol* **77**, 5721-5730 (2003).
913 21. Josefsson, L., *et al.* The HIV-1 reservoir in eight patients on long-term suppressive
914 antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci U*
915 *S A* **110**, E4987-4996 (2013).
916 22. Kearney, M.F., *et al.* Lack of detectable HIV-1 molecular evolution during suppressive
917 antiretroviral therapy. *PLoS Pathog* **10**, e1004010 (2014).
918 23. Nickle, D.C., *et al.* Evolutionary indicators of human immunodeficiency virus type 1
919 reservoirs and compartments. *J Virol* **77**, 5540-5546 (2003).
920 24. Rothenberger, M.K., *et al.* Large number of rebounding/founder HIV variants emerge
921 from multifocal infection in lymphatic tissues after treatment interruption. *Proc Natl Acad*
922 *Sci U S A* **112**, E1126-1134 (2015).
923 25. Brodin, J., *et al.* Establishment and stability of the latent HIV-1 DNA reservoir. *Elife*
924 **5**(2016).
925 26. Bull, M.E., *et al.* Monotypic human immunodeficiency virus type 1 genotypes across the
926 uterine cervix and in blood suggest proliferation of cells with provirus. *J Virol* **83**, 6020-
927 6028 (2009).
928 27. von Stockenstrom, S., *et al.* Longitudinal Genetic Characterization Reveals That Cell
929 Proliferation Maintains a Persistent HIV Type 1 DNA Pool During Effective HIV Therapy. *J*
930 *Infect Dis* **212**, 596-607 (2015).
931 28. Wagner, T.A., *et al.* An increasing proportion of monotypic HIV-1 DNA sequences during
932 antiretroviral treatment suggests proliferation of HIV-infected cells. *J Virol* **87**, 1770-1778
933 (2013).
934 29. Alizon, S. & Fraser, C. Within-host and between-host evolutionary rates across the HIV-1
935 genome. *Retrovirology* **10**, 49 (2013).
936 30. Hosmane, N.N., *et al.* Proliferation of latently infected CD4+ T cells carrying replication-
937 competent HIV-1: Potential role in latent reservoir dynamics. *J Exp Med* **214**, 959-972
938 (2017).
939 31. Maldarelli, F., *et al.* HIV latency. Specific HIV integration sites are linked to clonal
940 expansion and persistence of infected cells. *Science* **345**, 179-183 (2014).
941 32. Wagner, T.A., *et al.* HIV latency. Proliferation of cells with HIV integrated into cancer genes
942 contributes to persistent infection. *Science* **345**, 570-573 (2014).
943 33. Boritz, E.A., *et al.* Multiple Origins of Virus Persistence during Natural Control of HIV
944 Infection. *Cell* **166**, 1004-1015 (2016).
945 34. Cohn, L.B., *et al.* HIV-1 integration landscape during latent and active infection. *Cell* **160**,
946 420-432 (2015).
947 35. Simonetti, F.R., *et al.* Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo.
948 *Proc Natl Acad Sci U S A* **113**, 1883-1888 (2016).
949 36. Macallan, D.C., *et al.* Rapid turnover of effector-memory CD4(+) T cells in healthy humans.
950 *J Exp Med* **200**, 255-260 (2004).

951 37. McCune, J.M., *et al.* Factors influencing T-cell turnover in HIV-1-seropositive patients. *J*
952 *Clin Invest* **105**, R1-8 (2000).
953 38. Josefsson, L., *et al.* Single cell analysis of lymph node tissue from HIV-1 infected patients
954 reveals that the majority of CD4+ T-cells contain one HIV-1 DNA molecule. *PLoS Pathog* **9**,
955 e1003432 (2013).
956 39. Bruner, K.M., *et al.* Defective proviruses rapidly accumulate during acute HIV-1 infection.
957 *Nat Med* **22**, 1043-1049 (2016).
958 40. Ho, Y.C., *et al.* Replication-competent noninduced proviruses in the latent reservoir
959 increase barrier to HIV-1 cure. *Cell* **155**, 540-551 (2013).
960 41. Eren, M.I., Chao, A., Hwang, W.H. & Colwell, R.K. Estimating the richness of a population
961 when the maximum number of classes is fixed: a nonparametric solution to an
962 archaeological problem. *PLoS One* **7**, e34179 (2012).
963 42. Cardozo, E.F., *et al.* Treatment with integrase inhibitor suggests a new interpretation of
964 HIV RNA decay curves that reveals a subset of cells with slow integration. *PLoS Pathog* **13**,
965 e1006478 (2017).
966 43. Hill, A.L., Rosenbloom, D.I., Fu, F., Nowak, M.A. & Siliciano, R.F. Predicting the outcomes
967 of treatment to eradicate the latent reservoir for HIV-1. *Proc Natl Acad Sci U S A* **111**,
968 13475-13480 (2014).
969 44. Hunt, P.W., *et al.* T cell activation is associated with lower CD4+ T cell gains in human
970 immunodeficiency virus-infected patients with sustained viral suppression during
971 antiretroviral therapy. *J Infect Dis* **187**, 1534-1543 (2003).
972 45. Kaufmann, G.R., *et al.* Rapid restoration of CD4 T cell subsets in subjects receiving
973 antiretroviral therapy during primary HIV-1 infection. *AIDS* **14**, 2643-2651 (2000).
974 46. Perelson, A.S., *et al.* Decay characteristics of HIV-1-infected compartments during
975 combination therapy. *Nature* **387**, 188-191 (1997).
976 47. Palmer, S., *et al.* Low-level viremia persists for at least 7 years in patients on suppressive
977 antiretroviral therapy. *Proc Natl Acad Sci U S A* **105**, 3879-3884 (2008).
978 48. Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M. & Ho, D.D. HIV-1 dynamics
979 in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*
980 **271**, 1582-1586 (1996).
981 49. Reeves, D.B., *et al.* Anti-proliferative therapy for HIV cure: a compound interest approach.
982 *Sci Rep* **7**, 4011 (2017).
983 50. Shao, W., *et al.* Retrovirus Integration Database (RID): a public database for retroviral
984 insertion sites into host genomes. *Retrovirology* **13**, 47 (2016).
985 51. Chao, A. Estimating the population size for capture-recapture data with unequal
986 catchability. *Biometrics* **43**, 783-791 (1987).
987 52. Willis, A. Extrapolating abundance curves has no predictive power for estimating
988 microbial biodiversity. *Proc Natl Acad Sci U S A* **113**, E5096 (2016).
989 53. Perelson, A.S., Kirschner, D.E. & De Boer, R. Dynamics of HIV infection of CD4+ T cells.
990 *Math Biosci* **114**, 81-125 (1993).
991 54. Rutishauser, R.L., *et al.* Early and Delayed Antiretroviral Therapy Results in Comparable
992 Reductions in CD8+ T Cell Exhaustion Marker Expression. *AIDS Res Hum Retroviruses*
993 (2017).

994 55. Serrano-Villar, S.*, et al.* HIV-infected individuals with low CD4/CD8 ratio despite effective
995 antiretroviral therapy exhibit altered T cell subsets, heightened CD8+ T cell activation, and
996 increased risk of non-AIDS morbidity and mortality. *PLoS Pathog* **10**, e1004078 (2014).
997 56. Cockerham, L.R.*, et al.* Programmed death-1 expression on CD4(+) and CD8(+) T cells in
998 treated and untreated HIV disease. *AIDS* **28**, 1749-1758 (2014).
999 57. Appay, V. & Sauce, D. Immune activation and inflammation in HIV-1 infection: causes and
1000 consequences. *J Pathol* **214**, 231-241 (2008).
1001 58. Autran, B.*, et al.* Positive effects of combined antiretroviral therapy on CD4+ T cell
1002 homeostasis and function in advanced HIV disease. *Science* **277**, 112-116 (1997).
1003 59. Sanchez, M.A. & Blower, S.M. Uncertainty and sensitivity analysis of the basic
1004 reproductive rate. Tuberculosis as an example. *Am J Epidemiol* **145**, 1127-1137 (1997).
1005