Blind Uneven Proliferation of CD4+ T cells During Primary Infection Generates the Majority of the HIV Reservoir

Florencia A. T. Boshier^{a,1}, Daniel B. Reeves^a, Elizabeth R. Duke^{a, b}, David A. Swan^a, Martin Prlic^{a, c}, E. Fabian Cardozo-Ojeda^a, and Joshua T. Schiffer^{a, b, d}

^aVaccine and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave., Seattle, WA, 98122, USA

^bDepartment of Medicine, University of Washington, 1959 NE Pacific St., Seattle, WA, 98195, USA

^cDepartment of Global Health, University of Washington, 1959 NE Pacific St., Seattle, WA, 98195, USA

^dClinical Research Division, University of Washington, 1959 NE Pacific St., Seattle, WA.

April 6, 2020

Abstract

The HIV reservoir is a population of 1-10 million anatomically dispersed, latently infected memory CD4+ T cells in which an HIV DNA molecule is quiescently integrated into human chromosomal DNA. When antiretroviral therapy (ART) is stopped and HIV replication initiates in one of these cells, systemic viral spread resumes, rekindling progression to AIDS. Therefore, HIV latency prevents cure. The HIV reservoir contains clones: identical HIV sequences that are integrated within identical human chromosomal DNA locations. The presence of these clones demonstrates that proliferation of CD4+ T cells sustains infection despite ART. The reservoir has a precise structure consisting of a small number of large clones and a large number of small clones. However, the mechanisms leading to this structure have not been identified. We developed a mathematical model that recapitulates the profound depletion and brisk recovery of CD4+ T cells, reservoir creation, and viral load trajectory during primary HIV infection. We extended the model to simulate stochastically individual HIV reservoir clones and identified that uneven proliferation among clones during recovery from CD4+ lymphopaenia is sufficient to explain the observed clonal reservoir distribution. We project that within one month of infection 75-95% of reservoir cells are generated from cellular proliferation rather than denovo viral infection. Recent detection of HIV infected clones during the first 5 weeks of infection support our model's predictions.

The HIV latent reservoir is a population of resting memory CD4+ T cells that contain HIV-1 DNA integrated within their human chromosomal DNA [40, 27, 26, 16]. Resumption of HIV replication within one latently infected cell can spark systemic spread of virus, usually within days to weeks of antiretroviral therapy (ART) analytical treatment interruption (ATI) [32, 20, 39, 78]. The HIV reservoir size is stable over decades of ART [70, 23] and imperturbable, even if pharmacologically powerful agents are given that reactivate the virus in cell culture [69, 8]. Persistence of HIV latency therefore precludes cure and necessitates lifelong ART.

The HIV reservoir is established swiftly after initial infection. In non-human primate models, reservoir seeding occurs within three days of viral challenge [77, 76]. In humans, initiation of ART

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

within 10 days after estimated HIV infection leads to a reduced latent reservoir size [6] and substantial delays in viral rebound after ATI [41, 30, 28, 22] suggesting that the reservoir formation begins in the first week after HIV acquisition [31, 5].

Long-term studies of individuals on prolonged ART show that the reservoir often contains multiple cells carrying HIV DNA proviruses with identical sequences [36, 18, 48, 54, 15, 73, 72], which are integrated into the exact same human chromosomal integration site [73, 45, 25, 19]. Each unique sequence is believed to be generated from a distinct cellular infection event, as the HIV reverse transcriptase enzyme is highly error prone [79]. Identical viral sequences with matching integration sites suggest that these genomes were copied by the cellular DNA polymerase during the S phase of interphase, prior to mitosis. Consequently, the existence of abundant HIV DNA sequence clones proves that proliferation of infected memory CD4+ T cells helps sustain the HIV reservoir [59].

The clonal structure of the HIV reservoir appears to be highly organised. The rank according to size of an individual HIV cellular clone is likely to be predictive of the size of that clone. The reservoir is in turn defined by a small number of high-rank, large clones and a far larger number of low-rank, small clones [59]. This trend is evident when all HIV DNA is measured or when replication competent viruses are assessed alone [37]. A similar structure is observed for clones within the entire T cell repertoire as defined by T cell receptor (TCR) sequencing [42, 24, 55], which suggests that HIV reservoir clone dynamics may be partially or fully governed by more general rules of CD4+ T cell immunity [67]. However, the exact mechanisms and cellular dynamics leading to the precise rank-order distribution of the HIV reservoir are not known.

The HIV reservoir's clonal structure is clearly evident after a year or less of ART initiation and has been observed in individuals who started ART soon after primary infection [37, 34, 45, 73], suggesting that viral and cellular dynamics during early infection contribute to its formation. Primary HIV infection is characterised by profound CD4+ T cell lymphopaenia and peak viremia two weeks post-infection, followed by an incomplete, but brisk recovery in CD4+ T cell levels over the next four weeks. Stable HIV DNA and RNA setpoints are established concurrently approximately four weeks post-infection [5, 53, 63].

Formation of the HIV reservoir may be impacted by several coupled, non-linear processes including a mounting immune response [53], CD4+ T cell activation [50], and target cell limitation [61], necessitating the use of mathematical models to decipher available data. In previous modelling studies, generation of latently infected cells during untreated primary infection was ascribed entirely to viral infection of memory CD4+ T cells with subsequent conversion to a resting state [70, 9, 32, 60].

CD4+ T cell reconstitution is likely to be driven by homeostatic proliferation. Mouse models of CD4+ T cell recovery following depletion conclude that homeostatic proliferation is blind at the cell subset level with overall competition between CD4+ and CD8+ T cells for limited resources and space during homeostatic proliferation [49, 2, 3, 64, 44] but highly uneven at the single clone level with some TCR-defined clones outcompeting others due to variable binding affinity to self-antigens [3].

Here we develop a model of HIV primary infection that includes blind and uneven homeostatic proliferation of latently infected cells during recovery from CD4+ lymphopaenia. We fit the model to empirical data from 12 untreated individuals during primary infection [43, 52]. The model output suggests that homeostatic proliferation of CD4+ T cells drives the observed size and clonal structure of the HIV reservoir during the first several weeks of infection. It makes the empirically confirmed prediction that HIV sequence clones should be discoverable within a month of HIV infection despite the fact that most circulating HIV DNA sequences are in actively infected cells with unique non-clonal sequences [17].

1 Methods & Materials

1.1 Deterministic mathematical model

We developed several competing mathematical models to study the potential role of homeostatic proliferation in the formation of the latent reservoir during primary infection. All candidate models are derived from the following master system:

$$\begin{aligned}\frac{dE}{dt} &= r_E \left(1 - \frac{N + S + A + L + E}{\hat{K}} \right) E + w_1 h(A, E) \\ \frac{dN}{dt} &= r_N \left(1 - \frac{N + S + A + L + E}{\hat{K}} \right) N - \alpha V N + \theta \\ \frac{dS}{dt} &= r_S \left(1 - \frac{N + S + A + L + E}{\hat{K}} \right) S \\ &\quad - \beta V S + g \alpha V N \\ \frac{dL}{dt} &= r_L \left(1 - \frac{N + S + A + L + E}{\hat{K}} \right) L + f \beta V S - a L \\ \frac{dA}{dt} &= (1 - f) \beta V S - d_A A + a L - w_2 h(A, E) \\ \frac{dV}{dt} &= p A - c V\end{aligned}\tag{1}$$

Where $E(t)$ represents the concentration [cells/ μ L] of CD8+ T cells at time t . The CD4+ T cells are divided into four compartments: non-susceptible, susceptible, latently infected and actively infected. We denote the concentration [cells/ μ L] of these at time t by $N(t)$, $S(t)$, $L(t)$ and $A(t)$ respectively. For data fitting purposes, the reservoir of latently infected cells $L(t)$ represents the total HIV DNA rather than the much smaller subset of genomically intact or replication competent sequences [34]. Finally, $V(t)$ represents the plasma viral load over time [copies/ μ L]. We use the term all infected cells to refer to the sum of actively and latently infected cells.

The viral component of the model builds upon the canonical model of viral dynamics by differentiating activated and non-activated CD4+ T cells [33, 62, 75]. The virus is cleared at a rate c . Actively infected cells produce virus at a rate p , which describes the aggregate rate of constant viral leakage and burst upon cell death, and die at a rate d_A . The susceptible cells, S , are infected by the virus at a rate β [μ L/virus-day]. We assume that a fraction, f , of infection events generate latently infected cells and the remaining fraction, $(1 - f)$, leads to productively infected cells. Latently infected cells reactivate and become productively infected at a rate a [cell/ μ L-day].

In contrast to previous infection models, we incorporate the blind homeostasis hypothesis to explain cellular dynamics. We assume that there is an inherent competition for resources between CD8+ and CD4+ T cells that impact expansion and cell survival. We model this by logistically limiting the growth of the CD8+ T cells and the non-actively infected CD4+ T cell subsets by their sum. A similar mathematical formulation of blind homeostasis was considered in [49] to explain experiments of T cell reconstitution following CD4+ depletion in mice [3]. The turnover rates [cells/ μ L-day] are denoted by r_* . The functional carrying capacity of the system is denoted by K [cells/ μ L]. We considered a set of models that did not include this assumption by excluding the CD8+ T cell compartment entirely.

We evaluated the importance of several additional mechanisms. First, the killing of actively infected cells by CD8+ T cells at a rate w_2 , with or without the HIV antigen stimulation of CD8+ T cells at a rate w_1 as described by function $h(A, E)$. We consider two forms for $h(A, E)$, either linear or concentration-dependent as given by equation (2).

$$h(A, E) = AE \quad \text{or} \quad \frac{E}{E + E_{50}} A.\tag{2}$$

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Second, we considered the possibility that the virus removes non-susceptible cells from their compartment at a rate α [$\mu\text{L}/\text{virus-day}$]. A proportion $(1 - g)$ undergo bystander killing [47] and the remaining fraction become susceptible, i.e. upregulate CCR5, as they are part of the mounting cellular immune response to the virus. Finally, we investigated the possible role of generation of non-susceptible cells by the thymus at a constant rate θ . We assessed the above assumptions individually and in combination. Overall, we constructed 12 competing models. A detailed account of the model selection can be found in *Supporting Information S1*.

1.2 Stochastic mathematical model of individual clones within the latent reservoir

We are principally concerned with whether and how the proliferation of latently infected cells affects formation and sustainment of the clonal structure of the HIV reservoir during primary infection. To study the clonal distribution of the latent reservoir, we use a stochastic model in-place of equation 7. We let $L(t) = \sum_{n=1}^N L_n(t)$, where $L_n(t)$ denotes the number of cell members in clone n at time t and N denotes the number of distinct clones. In any short time-interval, four events can occur: a new distinct sequence can be introduced into the latent reservoir, an existing member of a clone can proliferate, a clone can contract in size or extinguish, or an existing member of a clone can convert to an actively infected cell and reactivate HIV.

Each new viral infection of a cell is associated with a unique HIV DNA sequence due to the error prone HIV reverse transcriptase, as well as a unique human chromosomal integration site [45, 54, 65, 66]. Therefore we assume that each infection event that generates a latently infected cell introduces a new viral sequence. We assume each new sequence population to have an initial size of 1, such that

$$N \rightarrow N + 1 \quad \text{and} \quad L_N : 0 \rightarrow 1 \quad (3)$$

with probability $f\beta V(t)S(t)$.

Each clone in the reservoir may die out stochastically or expand due to cellular proliferation. We refer to a clone with a single member, $L_n = 1$, as a singleton. Upon proliferation, a single HIV DNA molecule is copied with the human DNA polymerase, generating equivalent HIV DNA sequences and human chromosomal integration sites in the two daughter cells. Therefore we allow each clone in the reservoir to grow ($L_n \rightarrow L_n + 1$) with probability $r_{L,n}L_n/K$ and decline ($L_n \rightarrow L_n - 1$) with probability $(r_{L,n}(N + S + I + L + E)/K + a)L_n$. When $L_n > 1$ its HIV DNA sequence is considered to be part of a true clone in our model output.

The parameter $r_{L,n}$ represents the proliferation rate of the clone n . Based on the observations that distinct T cell clones proliferate at vastly different rates in response to lymphopaenia [3], and that the HIV reservoir is characterised by variability in clone size over 5-6 orders of magnitude [59], we allowed each clone to take on a different proliferation rate. Specifically, we defined uneven proliferation by assuming that the turnover rates are lognormally distributed, $r_{L,n} \sim \text{lognormal}(\mu, \sigma)$ with corresponding mean $m = \exp(\mu + \sigma^2)$ and variance $v = (\exp(\sigma^2) - 1) \exp(2\mu + \sigma^2)$.

Our simulations were run in Matlab. We simulated the deterministic compartments using ode45 [68] and used a Gillespie tau-leap of the same time-step for the stochastic compartment [29]. Each was updated simultaneously. Simulations were run in a litre of blood to simulate observed clinical values and to avoid numerical intractability of whole body simulations.

1.3 Parameter estimation

We fit our deterministic primary infection model (instances of equation (1)) to digitally extracted viral loads and matching CD4+ counts from 12 participants from the Females Rising through Education, Support and Health cohort (FRESH) [43, 52]. Our goal was to capture generalised trends in viral expansion, peak, contraction and setpoint, and CD4+ T cell contraction, re-expansion and

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

re-establishment of a setpoint. We employed a nonlinear, mixed-effects modelling approach implemented in Monolix [7]. Full details of model fitting can be found in *Supplementary information S1*. We fixed 9 of 17 parameters from reported values in the literature and estimated the remaining 8. For fitting the deterministic model we made a simplifying assumption that $r_L = r_N$.

We also fixed the initial conditions of the different variables in our model. We set $A(0) = 1$ and $L(0) = V(0) = 0$. As we have no CD8+ T cell data, we set $E(0) = 500$, the mean pre-infection values reported in the RV217 study [63], for models requiring it. We assume that susceptible cells are represented primarily by CCR5+ T cells, which in this cohort are found to be between 1-12 % of total CD4+ T cells in the plasma pre-infection [52]. We therefore set $S(0)$ to be 10% and $N(0)$ to be 90% of total CD4+ T cells for all participants. The resulting carrying capacity is found by assuming steady state pre-infection such that $K = E(0) + S(0) + N(0)$.

The parameters μ and σ governing clone proliferation rates and the latency fraction f are unknown. Therefore we performed sensitivity analyses to examine the effects of these three parameters on the size of the latent reservoir throughout infection and its clonal structure at 50 days post-infection.

2 Results

2.1 Blind proliferation between T cells explain CD4+ reconstitution during primary infection

The most parsimonious model by Akaike Information Criterion (AIC) score is shown in figure 1 and mathematically here:

$$\frac{dE}{dt} = r_E \left(1 - \frac{N + S + A + L + E}{K} \right) E \quad (4)$$

$$\frac{dN}{dt} = r_N \left(1 - \frac{N + S + A + L + E}{K} \right) N - \alpha V N \quad (5)$$

$$\frac{dS}{dt} = r_S \left(1 - \frac{N + S + A + L + E}{K} \right) S + g\alpha V N - \beta V S \quad (6)$$

$$\frac{dL}{dt} = r_L \left(1 - \frac{N + S + A + L + E}{K} \right) L + f\beta V S - aL \quad (7)$$

$$\frac{dA}{dt} = (1 - f)\beta V S - dA + aL \quad (8)$$

$$\frac{dV}{dt} = pA - cV. \quad (9)$$

Key assumptions required for model fit to data include blind homeostatic proliferation of both CD4+ and CD8+ T cells, density-dependent virus induced bystander killing of non-susceptible CD4+ T cells, and density-dependent virus induced activation of non-susceptible CD4+ T cells to susceptible CD4+ T cells via upregulation of CCR5+. CD8+ T cell expansion is driven solely by blind homeostatic proliferation. As studies in humans with hyperacute infection have shown that 5-20 % of activated CD8+ T cells during primary infection are responding to the virus [52] we also simulated an alternative model that incorporates HIV-induced CD8+ T cell activation. This model also fits the data but had a limited impact on latent reservoir dynamics as shown in figure S1.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Figure 1: Most parsimonious mathematical model of primary HIV infection. E represents CD8+ T cells. The CD4+ T cells are divided into four compartments: non-susceptible (N), susceptible (S), latently infected (L) and actively infected (A). The virus is represented by V .

The population level parameters found under these assumptions are shown in table 1. Certain solved model parameters correlate, hence we do not claim to have identified their absolute values. The individual parameter values used for each participant can be found in Supplementary table S2.

In figure 2, we show the model fit to data from a single study participant. For measurements of cellular concentrations, we scale simulation results, which are run in a liter of blood, to a μL and state viral load results in mL . Simulation results from our model closely recapitulate longitudinal measures of viral load (figure 2a) and total CD4+ counts, which is the sum of $N(t)$, $S(t)$, $L(t)$ and $A(t)$, (figure 2b). The expansion behaviour of the CD8+ T cell trajectories (figure 2c) while not available to us from this study, falls within the range observed during primary infection in other studies [63]. The reservoir volume (measured as total reservoir HIV DNA in figure 2d) expands to levels observed during chronic ART treatment [14] with an HIV DNA setpoint within 2 weeks of infection. Model fits for the other study participants are shown in figure S1. Predicted trajectories for $N(t)$, $S(t)$, $L(t)$ and $A(t)$ suggest that each of these cell populations equilibrate to a quasi-steady state 20-40 days after infection.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Figure 2: Longitudinal results of deterministic mathematical model fits (black line) to data from participant 11 (dots) for: a) viral load b) CD4+ T cells = $N(t) + S(t) + A(t) + L(t)$, c) CD8+ T cells d) Latent reservoir size. X-axis shown in days following onset plasma viraemia (DOFPV).

2.2 Homeostatic cellular proliferation is required to recapitulate the inverse correlation observed between CD4+ T cell nadir and HIV reservoir volume

Previous studies identified that levels of HIV-1 proviral DNA during chronic ART are negatively correlated with CD4+ T cell nadir during primary infection [11] particularly if ART is started during acute infection [71]. In order to establish whether this correlation occurs during untreated primary infection in our simulations, we compare CD4+ T cell nadir and size of the latent reservoir at 30 days post-infection under our deterministic model as shown in figure 3. Simulating all participants using parameter values derived from our individual fits, table S2, we identify that CD4 T-cell nadir as estimated from the model had a significant, negative correlation with size of the latent reservoir 30 days post-infection ($r^2 = 0.34$, $p < 0.05$). At later times during infection, the correlation weakens in agreement with [71]. However no correlation exists if reservoir proliferation is excluded (i.e. set $r_L = 0$), which suggests that the observed correlation is mechanistically related to latent cell proliferation during primary infection.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Table 1: Summary of deterministic model parameters and initial conditions. When estimated, the value shown is from the population level fits

Parameter	Value	Dimension	Source
t_0	-3.76	days	estimated
E_0	500	cells/ μ L	[63]
N_0	$0.90 \times T_0$	cells/ μ L	fixed
S_0	$0.10 \times T_0$	cells/ μ L	fixed
K	$E_0 + N_0 + S_0$	cells/ μ L	fixed
r_E	0.085	cells/(μ L-day)	estimated
r_N	0.251	cells/(μ L-day)	estimated
r_S	0.311	cells/(μ L-day)	estimated
r_L	0.251	cells/(μ L-day)	assumed equal to r_N
α	1.2×10^{-5}	μ L/(virus-day)	estimated
g	0.26	NA	estimated
f	10^{-4}	NA	[21]
β	$10^{-4.01}$	μ L/(virus-day)	estimated
d	1	day $^{-1}$	[46]
p	10^4	virus/(cell-day)	[56, 35]
c	23	day $^{-1}$	[58]
a	5.75×10^{-5}	day $^{-1}$	[4]

2.3 Individual clone expansion is predicted to occur immediately after reservoir formation

Next we simulated our stochastic model with individual clones. A single realization of the latent reservoir is shown in figure 4. Expanding clones are observed in the latent reservoir less than a day after it is seeded, and new HIV DNA sequences are added to the reservoir continuously throughout untreated infection due to new infection events. The individual clones sum to a reasonable approximation of reservoir size, comparable to the deterministic model simulations in figure 2d.

Assuming otherwise unchanged infection dynamics, HIV reservoir size could potentially be determined by three parameters in our stochastic model: the latency fraction, f (fraction of newly infected cell entering latency), mean clone proliferation rate, m , and variance of proliferation rates across clones, v . We find that altering the latency fraction by several orders of magnitude has a negligible impact on the projected size of the reservoir (figure S2). Increasing latent cell proliferation rates to twice those in other non-susceptible CD4 T cells results in several-fold larger reservoir volumes (figure S2a, S3a). Increasing the variance in proliferation rate across reservoir clones also generates larger reservoir volumes (figure S2b, S3b) and increases the sensitivity of the model to latency fraction in a non-linear fashion (figure S2b). Whilst moderate mean proliferation rates and variances lead to plausible latent reservoir sizes in stochastic model simulations, very large mean rates and/or variances may result in unrealistically large latent reservoirs comprising more than 50% of an individual's CD4+ T cell repertoire (figure S3). Therefore we are able to establish upper bounds on values for the mean and variance of latent cell proliferation rate.

2.4 The latent reservoir reaches its highest sequence richness during the first 10 days of HIV infection

Next we consider the dynamics of sequence richness of the latent reservoir. Sequence richness is the number of unique total HIV DNA sequences or chromosomal integration sites in the body and is a

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Figure 3: Inverse correlation between CD4+ nadir and latent reservoir volume at 30 days after primary infection in model simulations. a) An inverse correlation between the latent reservoir and CD4 + T cell nadir exists if the latent reservoir proliferates ($r_L \neq 0$) ($p < 0.05$). Each colour corresponds to one of twelve participants in the FRESH cohort. b) Correlation between the latent reservoir and CD4 + T cell nadir disappears if the latent reservoir does not to proliferate ($r_L = 0$).

marker of the number of latent cells generated by de novo HIV infection. Mathematically, we define sequence richness at time t , $R(t)$, as

$$R(t) = \|(L_1(t), L_2(t), \dots, L_N(t))\|_0. \quad (10)$$

As shown in figure 5a, in model realizations of all study participants, the reservoir rapidly accumulates richness during the first ~ 10 days of infection due to unchecked viral replication leading to widespread infection of new cells with unique sequence and chromosomal integration sites. Following peak viraemia, the richness declines gradually towards a steady state as stochastic burnout of some clones balances formation of new sequences due to new ongoing infection events.

2.5 Proliferative clones predominate in the latent reservoir within two weeks of primary HIV infection

In model simulations, expansion of sequence richness terminates at the CD4+ T cell nadir ~ 10 days post infection (figure 2b,d). This is followed by 2-4 weeks of rapid blind proliferation of CD4+ and CD8+ T cells to allow recovery from CD4+ lymphopaenia. As part of this process, homeostatic proliferation occurs among individual CD4+ T cell clones within the latent reservoir (figure 5b). Within two weeks of infection 12% to 71% of the latent reservoir is predicted to consist of clones derived from proliferation; within 100 days of infection 82% to 98% of the latent reservoir consists of clones derived from proliferation (figure 5c).

We normalised sequence richness according to the total number of latently infected cells (figure 5d). The maximum value of this ratio approaches 0.88-1.0 within the latent reservoir 10 days after infection as nearly all reservoir HIV DNA sequences are unique prior to clonal expansion. However, this ratio rapidly declines over the next 2 weeks to 0.5 by day 9-20 and 0.1 by day 18-88. This result implies that while viral infection establishes the reservoir extremely early after HIV trans-

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Figure 4: Simulated single clone dynamics during primary HIV infection. Each clone is represented by a single coloured line. The sum of all clones is shown in black. The mean of r_L , $m = r_N$ and the variance $v = 0.2r_N$.

mission, the reservoir is maintained primarily by clonal cellular proliferation in response to CD4+ lymphopaenia following peak viraemia.

2.6 Proliferative reservoir clones are obscured by a larger population of non-clonal sequences within actively infected cells during untreated infection

There are currently no experimental methods available to isolate latent reservoir HIV DNA sequences from sequences within cells harbouring active HIV replication during untreated infection. Instead, the latent reservoir is obscured by a population of rapidly mutating sequences from actively infected cells, each with a unique genetic signature and chromosomal integration site.

However, we can distinguish latently infected cells from actively infected cells in our model. Therefore we performed the above analysis inclusive of actively infected cells, each of which contributes a unique sequence singlet to the overall population of infected cells. Early during untreated infection, the number of actively infected cells outnumbers the number of latently infected cells by 1-3 orders of magnitude, such that observed richness is driven mostly by actively infected cells (figure 6a compared to figure 5a). During the first 10 days of infection, the proportion of clones among all sequences is extremely low (range: 10^{-5} - 6×10^{-4}) (figure 6b).

Upon establishment of a viral load and reservoir steady state at 25-50 days, most HIV DNA sequences obtained from a clinical sample are truly unique and the ratio of clones to population size is still relatively low (range: 5×10^{-3} - 3×10^{-1}) figure 6b, as actively infected cells still outnumber clonally expanding latently infected cells. Accordingly, the ratio of richness to the total population remains high figure 6c. Within 100 days of infection 2×10^{-3} to 7×10^{-1} of all infected cells are

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

predicted to consist of clones figure 6b.

We examined whether these predictions were robust to uncertainty regarding key reservoir model parameters. Lower mean proliferation and higher variance in proliferation rate results in slight delays in the timing of clonal predominance within the HIV reservoir as does higher variance in turnover rates (figure S4a,b,c,d). While these parameters may indeed vary across infected persons, it is a generalisable property of our model output that a majority of the latent reservoir (figure S4a,c), but a minority of all infected cells (figure S4b,d) are composed of sequence clones within a month of infection.

Figure 5: Establishment of clonal predominance in the HIV reservoir within the first month of infection due to rapid homeostatic proliferation of CD4+ T cells from stochastic model simulations. (a) Richness of HIV DNA sequences within latently infected reservoir cells. (b) Number of HIV DNA reservoir sequences within proliferative CD4+ T cell clones. (c) Clonal sequence : total sequence population size ratio for latently infected cells. (d) Richness : total sequence population size ratio for latently infected cells. Each colour corresponds to an individual as in figure 3.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Figure 6: HIV DNA reservoir sequences are obscured by a cloud of non-clonal sequences from actively infected cells during early untreated infection in stochastic model simulations. (a) Richness of HIV DNA sequences within all infected cells. (b) Clonal sequence : total sequence population size ratio within all infected cells. (c) Richness : total sequence population size ratio within all infected cells. Each colour corresponds to an individual as in figure 3.

2.7 Adequate sequence sample size ensures detection of HIV DNA clones within 2-3 weeks of primary infection

Our model projects that by Fiebig stage IV (day 27), 70-95 % of HIV DNA sequences within the reservoir are clonal (figure 5 c). However, only 2-20% of total HIV DNA sequences are clonal (figure 6b) making the detection of clones in the reservoir challenging. Another barrier to detecting clonal HIV DNA from the reservoir is that the sequence sample size in most reservoir studies are low, varying between ~ 10 in studies studying intact, replication competent viruses and ~ 100 -1000 in studies of total HIV DNA [37, 34, 45, 73]. Therefore, in studies of ART treated study participants, many observed singleton sequences may actually be members of large proliferative clones [59].

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

2.8 Predictions are in keeping with recent study

In a recent study of untreated individuals, clones were not observed in 3 participants sampled during Fiebig III-IV stages of infection (days 22-27); but 4 participants sampled during Fiebig stage IV-V (~day 30-70 post infection) had at least one detectable HIV DNA clone based on chromosomal integration site analysis; [17]. The number of HIV DNA integration sites sampled in the Fiebig IV-V participants varied from 48 - 709. The mean percentage of clones among all sequences in the Fiebig stage V participants was extremely low (0.72%).

Results from this study are in keeping with our model's predictions (figure 7). We performed sampling of HIV DNA sequences from model-generated data to identify sequence sample size required for detection of clones at different times post-infection (corresponding to distinct Fiebig stages). The fraction of the sample identified as clonal is shown in figure 7.

If technology existed to sample only HIV DNA sequences from the latently infected reservoir, then sampling of 1000 sequences would reliably detect clones by Fiebig I while sampling of 100 sequences would usually detect clones by Fiebig III (figure 7, left column). However, with sampling of all HIV DNA sequences (figure 7, right column), clones would almost never be evident during Fiebig stages I-III. A sequence sample size of 1000 would sometimes allow detection of a low percentage of clones during Fiebig IV as in [17] and usually allow their detection by Fiebig stage V. A sequence sample size of 100 would reliably detect clones only after 2 months of untreated infection (Fiebig V) (figure 7).

These results are dependent on the mean and variance of turnover rates. Lower mean (figure S5a) and variance (figure S5b) of latent cell proliferation rates results in a lower probability of detection of clones 50 days after infection. Nevertheless, if any latent reservoir proliferation occurs our model predicts that proliferative clones should be possible from Fiebig Stage III of active infection given a sequence sample size of 100 HIV DNA sequences.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Figure 7: Increased likelihood of detection of proliferative clones within the reservoir at later sampling times and greater HIV DNA sequence sample size from stochastic model simulations. Sampling from: a) Latently Infected Cells only, b) All Infected Cells. Each colour represents simulation results for a different number of cells sampled (10, 100, 1000). The numerical values in parentheses above the boxes indicate the mean percent of clonal cells.

2.9 Homeostatic proliferation establishes the clonal distribution of the latent reservoir within 2-3 weeks of primary infection

We demonstrated previously that the HIV reservoir has a fixed clonal structure consisting of a small number of large clones and a large number of smaller clones. The rank of each clone according to its size can be used to roughly estimate its size based on a power law relationship with a slope that approximates one [59].

In figure 8 we compare the clonal distribution of the latent reservoir within a single individual at different Fiebig stages. Clonal structure within the simulated reservoir is evident by Fiebig stage II ($\sim$ day 17), immediately after recovery from CD4⁺ T cell nadir (figure 8, orange). The uneven homeostatic proliferation of individual HIV DNA clones in the reservoir is sufficient to explain HIV reservoir structure. As infection progresses, the largest clones comprises a greater proportion of the reservoir (figure 8 a, left). This remains true even if we consider the indistinguishable HIV sequences arising from actively infected cells. A long tail of singleton sequences, clones containing only one member $L_n = 1$, is observed at all Fiebig stages during untreated infection (figure 8 a, right). The rank-order distributions are also somewhat steeper at later Fiebig stages as the richness of the reservoir decreases (figure 8a).

When we simulated the reservoir after one year of ART started at Fiebig Stage V (100 days post-infection), the clonal distribution of the simulated reservoir after 1 year of effective ART treatment is similar to that of the point at which ART was initiated, though the singleton tail is effectively eliminated. Notably, the rank-order distribution is equivalent between all infected cells and latently infected cells after a year of ART due to the elimination of actively infected cells (figure 8a light-blue).

Next we carried out an equivalent analysis assuming clinical sampling of 200 HIV DNA sequences (figure 8 b). Only at Fiebig stages IV and V were multiple clones detected with sampling of sequences from all infected cells, despite our prediction that the clonal structure is established 2 weeks earlier. The distributions we observed after one-year of ART resembled clonal distributions observed during ART in study participants [59], exhibiting a tail of observed singletons which are in fact likely to be members of larger clones.

Varying the mean proliferation rate of clones does not impact the rank-order slope (figure S6a) or the observed distribution of clone frequency assuming known sampling limitations (figure S6b). With no proliferation, our model projects that the reservoir would be composed entirely of singletons and the clonal distribution would not be distinct from a uniform distribution (Kolmogorov-Smirnoff test, $p > 0.05$). The clonal distribution contains a slightly greater proportion of large clones at a higher proliferation rate relative to a lower proliferation rate (Kolmogorov-Smirnoff test, $p < 0.05$). This distinction is lost with a sequence sample size of 200 (figure S6b).

The clonal distribution becomes steeper with increasing variance of the proliferation rates due to steeper rank abundance curves at higher variances (figure S6c) (Kolmogorov-Smirnoff test, $p = 0.05$). This distinction is maintained even with a sequence sample size of 200 (figure S6d).

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Figure 8: Rank abundance curves from stochastic model simulations for a single participant during sequential stages of early HIV infection. a) Rank abundance curves of all HIV DNA sequences in latently infected cells only (left) and all infected cells (right). b) Rank abundance curves of 200 randomly sampled HIV DNA sequences in latently infected cells only (left) and all infected cells (right).

3 Discussion

The primary obstacle to eradication of HIV is a reservoir of latently infected resting memory CD4+ T cells [70, 57]. Our mathematical model predicts that the HIV reservoir establishes a steady population during primary infection due to rapid, blind and uneven homeostatic CD4+ T cell proliferation. At the core of our assumptions is that HIV infected reservoir cells behave according to rules governing overall T cell dynamics during recovery from lymphopaenia [74].

Our model generates realistic trajectories for HIV viral load, CD4+ T cell contraction and re-expansion, and CD8+ T cell count expansion during primary infection. The model also recapitulates the reservoir's early formation, its ultimate size and its clonal distribution. Importantly, our model's

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

results are consistent with recent data showing that HIV DNA clones become detectable in approximately 50 % HIV infected individuals 1-2 months after primary infection [17].

Sampling the reservoir during early infection is difficult for two reasons. First, most observed HIV DNA sequences are from actively infected cells that have a much shorter lifespan than reservoir cells, harbour truly unique HIV DNA sequences and cannot be discriminated experimentally from reservoir-derived sequences [10]. Second, current studies dramatically under sample the reservoir: during prolonged ART, sequences which are only observed once are likely to be members of large clones rather than truly unique [59]. Despite these issues, our model makes the prediction that repeated HIV DNA integration sites can sometimes be detected within the first 1-2 months of HIV infection, consistent with the results in [17]. While doublets and triplets, clones with two or three members respectively, are only sporadically observed at these timepoints even with large sample sizes, we predict that by this timepoint, there is already a fully formed, mature HIV DNA reservoir with an organised clonal structure. In fact, the doublets and triplets are likely to represent extremely large clones whose size is obscured by under sampling, consisting of more than a million cells each. This signifies that certain clones may undergo roughly twenty divisions during the 2-3 weeks recovery period from CD4+ lymphopaenia.

While observed heterogeneity in clone size may be linked mechanistically with known properties of homeostatic proliferation, we cannot rule out the possibility that the largest clones may be generated due to recognition of a cognate antigen presented by HIV-infected cells or by cells infected with other ubiquitous chronic viruses such as cytomegalovirus or Epstein-Barr Virus.

Our conclusions demonstrate that cure interventions, even those given during primary infection, face an enormous challenge. Even limited seeding of the reservoir will be followed by a period of homeostatic proliferation during recovery from CD4+ nadir that will expand the size of the reservoir ~10-fold. Lymphocyte anti-proliferative agents have been proposed as a means to reduce reservoir size during chronic ART [60]. While this approach might be considered in an adjunctive fashion during primary infection, the immunosuppressive features of these drugs raise concerns that they may hamper the long-term cytolytic immune response to HIV. Based on comparable experience following stem cell transplantation [60], it is also questionable whether these agents could sufficiently suppress the powerful homeostatic forces underlying recovery from lymphopaenia.

Our model does reinforce the utility of ART during the first week of acute HIV in limiting reservoir volume [5]. Presumably, this intervention would limit the extent of reservoir seeding, and also limit the extent of CD4+ T cell nadir decline, thereby decreasing the extent of homeostatic proliferation required to reconstitute the overall CD4+ T cell population.

Several important limitations of our study are worth noting. First, while we establish that recovery from CD4+ lymphopaenia and uneven homeostatic proliferation are sufficient to generate the observed clonal structure of the HIV reservoir and argue that the reservoir is formed rapidly during primary infection, key reservoir-sustaining events continue to occur through all subsequent stages of infection. HIV DNA clones wax and wane in size during chronic ART, which suggests that substantial dynamics exists within the reservoir years after recovery from lymphopaenia [74]. These clonal expansions and contraction may relate to recognition of cognate antigen by certain clones rather than a homeostatic process [74].

Moreover, a majority of reservoir sequences observed during prolonged ART arise from timepoints immediately before ART initiation, suggesting that the biologic events which generate reservoir cells continue after primary infection [12, 1]. ART initiation is also followed by a surge in levels of activated, proliferating CD4+ T cells [38]. This process could also further expand a pre-existing reservoir. Neither of these processes is mutually exclusive of homeostatic proliferation as a driver of reservoir formation during primary HIV infection, and the latter has substantial parallels.

Second, in this study, we focus on total HIV DNA, but only a small fraction of these sequences is replication-competent [13]. As a result, the clones that have been discovered during early HIV infection are probably not replication competent. Nevertheless, viral rebound in patients who initiate ART after only days of infection proves that the replication-competent reservoir is created extremely

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

early during primary HIV [22, 41, 28]. Because replication competent HIV periodically reactivates and is targeted by the immune system, and because viral activation is unpredictably coupled with cellular proliferation [51], it remains unclear whether all conclusions drawn from modelling total HIV DNA during primary infection can be extrapolated to the replication competent reservoir.

Finally, our most parsimonious model does not include an immune response to the virus during primary infection. Our objective was not to identify different drivers of viral load set point or CD4+ T cell nadir, and thus we opted for a simpler model. In our model, the CD4+ and CD8+ T cell equation does not incorporate the possibility that as many as 5-20% of these cells may be reacting to HIV-infected cells and therefore have proliferation governed by HIV antigen abundance rather than homeostasis or expansion due to exposure to other non-HIV antigens [52]. However, we performed deterministic simulations including CD8+ T cell activation which suggest that this effect is unlikely to negate the critical role of homeostatic proliferation in rapidly restoring lymphocyte count following depletion.

In summary, we developed the hypothesis that blind, uneven homeostatic T cell recovery from CD4+ lymphopaenia drives the bulk of early HIV reservoir generation. Future curative therapies will need to address the mechanisms that drive the expansion behaviours of individual T cell clones.

References

- [1] M. R. Abrahams, S. B. Joseph, N. Garrett, L. Tyers, M. Moeser, N. Archin, O. D. Council, D. Matten, S. Zhou, D. Doolabh, C. Anthony, N. Goonetilleke, S. A. Karim, D. M. Margolis, S. K. Pond, C. Williamson, and R. Swanstrom. The replication-competent HIV-1 latent reservoir is primarily established near the time of therapy initiation. *Science Translational Medicine*, 11(513), oct 2019.
- [2] L. M. Adleman and D. Wofsy. T-cell homeostasis: implications in HIV infection. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 6:144–152, 1993.
- [3] L. M. Adleman and D. Wofsy. Blind T-cell homeostasis in CD4-deficient mice. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 11(4):334–340, 1996.
- [4] b. L. Alison Hill, D. I. S Rosenbloom, F. Fu, M. A. Nowak, and R. F. Siliciano. Correction for "Predicting the outcomes of treatment to eradicate the latent reservoir for. *Proc Natl Acad Sci*, 2014.
- [5] J. Ananworanich, N. Chomont, L. A. Eller, E. Kroon, S. Tovanabutra, M. Bose, M. Nau, J. L. Fletcher, S. Tipsuk, C. Vandergeeten, R. J. O'Connell, S. Pinyakorn, N. Michael, N. Phanuphak, and M. L. Robb. HIV DNA Set Point is Rapidly Established in Acute HIV Infection and Dramatically Reduced by Early ART. *EBioMedicine*, 11:68–72, sep 2016.
- [6] J. Ananworanich and M. L. Robb. The transient HIV remission in the Mississippi baby: why is this good news? *Journal of the International AIDS Society*, 17(1):19859, jan 2014.
- [7] . Antony, France: Lixsoft SAS. *Monolix version 2018R1*. <http://lixoft.com/products/monolix/>, 2018.
- [8] N. M. Archin, J. L. Kirchherr, J. A. Sung, G. Clutton, K. Sholtis, Y. Xu, B. Allard, E. Stuelke, A. D. Kashuba, J. D. Kuruc, J. Eron, C. L. Gay, N. Goonetilleke, and D. M. Margolis. Interval dosing with the HDAC inhibitor vorinostat effectively reverses HIV latency. *Journal of Clinical Investigation*, 127(8):3126–3135, aug 2017.
- [9] N. M. Archin, N. K. Vaidya, J. D. Kuruc, A. L. Liberty, A. Wiegand, M. F. Kearney, M. S. Cohen, J. M. Coffin, R. J. Bosch, C. L. Gay, J. J. Eron, D. M. Margolis, and A. S. Perelson.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

Immediate antiviral therapy appears to restrict resting CD4+ cell HIV-1 infection without accelerating the decay of latent infection. *Proceedings of the National Academy of Sciences*, 109(24):9523–9528, 2012.

- [10] J. N. Blankson, D. Finzi, T. C. Pierson, B. P. Sabundayo, K. Chadwick, J. B. Margolick, T. C. Quinn, and R. F. Siliciano. Biphasic Decay of Latently Infected CD4 + T Cells in Acute Human Immunodeficiency Virus Type 1 Infection . *The Journal of Infectious Diseases*, 182(6):1636–1642, dec 2000.
- [11] M. R. Boulassel, N. Chomont, N. P. Pai, N. Gilmore, R. P. Sékaly, and J. P. Routy. CD4 T cell nadir independently predicts the magnitude of the HIV reservoir after prolonged suppressive antiretroviral therapy. *Journal of Clinical Virology*, 53(1):29–32, 2012.
- [12] J. Brodin, F. Zanini, L. Thebo, C. Lanz, G. Bratt, R. A. Neher, and J. Albert. Establishment and stability of the latent HIV-1 DNA reservoir. *eLife*, 5, 2016.
- [13] K. M. Bruner, Z. Wang, F. R. Simonetti, A. M. Bender, K. J. Kwon, S. Sengupta, E. J. Fray, S. A. Beg, A. A. Antar, K. M. Jenike, L. N. Bertagnolli, A. A. Capoferri, J. T. Kufera, A. Timmons, C. Nobles, J. Gregg, N. Wada, Y. C. Ho, H. Zhang, J. B. Margolick, J. N. Blankson, S. G. Deeks, F. D. Bushman, J. D. Siliciano, G. M. Laird, and R. F. Siliciano. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses, feb 2019.
- [14] A. Cheret, C. Bacchus-Souffan, V. Avettand-Fenoel, A. Mélard, G. Nembot, C. Blanc, A. Samri, A. Sáez-Cirión, L. Hocqueloux, C. Lascoux-Combe, C. Allavena, C. Goujard, M. Valantin, A. Leplatôis, L. Meyer, C. Rouzioux, and B. Autran. Combined art started during acute hiv infection protects central memory cd4+ t cells and can induce remission. *The Journal of antimicrobial chemotherapy*, 70, 04 2015.
- [15] N. Chomont, M. El-Far, P. Ancuta, L. Trautmann, F. A. Procopio, B. Yassine-Diab, G. Boucher, M. R. Boulassel, G. Ghattas, J. M. Brenchley, T. W. Schacker, B. J. Hill, D. C. Douek, J. P. Routy, E. K. Haddad, and R. P. Sékaly. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature Medicine*, 15(8):893–900, 2009.
- [16] T. W. Chun, L. Carruth, D. Finzi, X. Shen, J. A. DiGiuseppe, H. Taylor, M. Hermankova, K. Chadwick, J. Margolick, T. C. Quinn, Y. H. Kuo, R. Brookmeyer, M. A. Zeiger, P. Barditch-Crovo, and R. F. Siliciano. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, 387(6629):183–8, may 1997.
- [17] J. M. Coffin, D. W. Wells, J. M. Zerbato, J. D. Kuruc, S. Guo, B. T. Luke, J. J. Eron, M. Bale, J. Spindler, F. R. Simonetti, S. Hill, M. F. Kearney, F. Maldarelli, X. Wu, J. W. Mellors, and S. H. Hughes. Clones of infected cells arise early in HIV-infected individuals. *JCI Insight*, 4(12):0–15, 2019.
- [18] L. B. Cohn, I. T. da Silva, R. Valieris, A. S. Huang, J. C. C. Lorenzi, Y. Z. Cohen, J. A. Pai, A. L. Butler, M. Caskey, M. Jankovic, and M. C. Nussenzweig. Clonal CD4+ T cells in the HIV-1 latent reservoir display a distinct gene profile upon reactivation. *Nature medicine*, 24(5):604–609, 2018.
- [19] L. B. Cohn, I. T. Silva, T. Y. Oliveira, R. A. Rosales, E. H. Parrish, G. H. Learn, B. H. Hahn, J. L. Czartoski, M. J. McElrath, C. Lehmann, F. Klein, M. Caskey, B. D. Walker, J. D. Siliciano, R. F. Siliciano, M. Jankovic, and M. C. Nussenzweig. HIV-1 integration landscape during latent and active infection. *Cell*, 160(3):420–432, jan 2015.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

- [20] D. J. Colby, L. Trautmann, S. Pinyakorn, L. Leyre, A. Pagliuzza, E. Kroon, M. Rolland, H. Takata, S. Buranapraditkun, J. Intasan, N. Chomchey, R. Muir, E. K. Haddad, S. To-vanabutra, S. Ubolyam, D. L. Bolton, B. A. Fullmer, R. J. Gorelick, L. Fox, T. A. Crow-ell, R. Trichavaroj, R. O’Connell, N. Chomont, J. H. Kim, N. L. Michael, M. L. Robb, N. Phanuphak, J. Ananworanich, and R. study group. Rapid HIV RNA rebound after an-tiretroviral treatment interruption in persons durably suppressed in Fiebig I acute HIV infec-tion. *Nature medicine*, 24(7):923–926, 2018.
- [21] J. M. Conway and A. S. Perelson. Residual Viremia in Treated HIV+Individuals. *PLoS Com-putational Biology*, 2016.
- [22] J. M. Conway, A. S. Perelson, and J. Z. Li. Predictions of time to HIV viral rebound following ART suspension that incorporate personal biomarkers. *PLoS Computational Biology*, 15(7), jul 2019.
- [23] A. M. Crooks, R. Bateson, A. B. Cope, N. P. Dahl, M. K. Griggs, J. A. D. Kuruc, C. L. Gay, J. J. Eron, D. M. Margolis, R. J. Bosch, and N. M. Archin. Precise quantitation of the latent HIV-1 reservoir: Implications for eradication strategies. *Journal of Infectious Diseases*, 212(9):1361–1365, nov 2015.
- [24] J. Desponds, T. Mora, and A. M. Walczak. Fluctuating fitness shapes the clone-size distribution of immune repertoires. *Proceedings of the National Academy of Sciences of the United States of America*, 113(2):274–279, jan 2016.
- [25] K. B. Einkauf, G. Q. Lee, C. Gao, R. Sharaf, X. Sun, S. Hua, S. M. Chen, C. Jiang, X. Lian, F. Z. Chowdhury, E. S. Rosenberg, T.-W. Chun, J. Z. Li, X. G. Yu, and M. Lichterfeld. Accumulation of intact proviruses at distinct chromosomal positions during long-term antiretroviral therapy. *Journal of Clinical Investigation*, in-press(3), 2019.
- [26] D. Finzi, J. Blankson, J. D. Siliciano, J. B. Margolick, K. Chadwick, T. Pierson, K. Smith, J. Lisiewicz, F. Lori, C. Flexner, T. C. Quinn, R. E. Chaisson, E. Rosenberg, B. Walker, S. Gange, J. Gallant, and R. F. Siliciano. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nature medicine*, 5(5):512–7, may 1999.
- [27] D. Finzi, M. Hermankova, T. Pierson, L. M. Carruth, C. Buck, R. E. Chaisson, T. C. Quinn, K. Chadwick, J. Margolick, R. Brookmeyer, J. Gallant, M. Markowitz, D. D. Ho, D. D. Rich-man, and R. F. Siliciano. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*, 278(5341):1295–1300, nov 1997.
- [28] P. Garcia-Broncano, S. Maddali, K. B. Einkauf, C. Jiang, C. Gao, J. Chevalier, F. Z. Chowdhury, K. Maswabi, G. Ajibola, S. Moyo, T. Mohammed, T. Ncube, J. Makhema, P. Jean-Philippe, X. G. Yu, K. M. Powis, S. Lockman, D. R. Kuritzkes, R. Shapiro, and M. Lichterfeld. Early antiretroviral therapy in neonates with HIV-1 infection restricts viral reservoir size and induces a distinct innate immune profile. *Science translational medicine*, 11(520), nov 2019.
- [29] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115(4):1716–1733, 2001.
- [30] T. J. Henrich, H. Hatano, O. Bacon, L. E. Hogan, R. Rutishauser, A. Hill, M. F. Kearney, E. M. Anderson, S. P. Buchbinder, S. E. Cohen, M. Abdel-Mohsen, C. W. Pohlmeier, R. Fromentin, R. Hoh, A. Y. Liu, J. M. McCune, J. Spindler, K. Metcalf-Pate, K. S. Hobbs, C. Thanh, E. A. Gibson, D. R. Kuritzkes, R. F. Siliciano, R. W. Price, D. D. Richman, N. Chomont, J. D. Siliciano, J. W. Mellors, S. A. Yukl, J. N. Blankson, T. Liegler, and S. G. Deeks. HIV-1 persistence following extremely early initiation of antiretroviral therapy (ART) during acute HIV-1 infection: An observational study. *PLoS Medicine*, 14(11), 2017.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

- [31] T. J. Henrich, H. Hatano, O. Bacon, L. E. Hogan, R. Rutishauser, A. Hill, M. F. Kearney, E. M. Anderson, S. P. Buchbinder, S. E. Cohen, M. Abdel-Mohsen, C. W. Pohlmeier, R. Fromentin, R. Hoh, A. Y. Liu, J. M. McCune, J. Spindler, K. Metcalf-Pate, K. S. Hobbs, C. Thanh, E. A. Gibson, D. R. Kuritzkes, R. F. Siliciano, R. W. Price, D. D. Richman, N. Chomont, J. D. Siliciano, J. W. Mellors, S. A. Yukl, J. N. Blankson, T. Liegler, and S. G. Deeks. HIV-1 persistence following extremely early initiation of antiretroviral therapy (ART) during acute HIV-1 infection: An observational study. *PLoS medicine*, 14(11):e1002417, nov 2017.
- [32] A. L. Hill, D. I. S. Rosenbloom, F. Fu, M. A. Nowak, and R. F. Siliciano. Predicting the outcomes of treatment to eradicate the latent reservoir for HIV-1. *Proceedings of the National Academy of Sciences*, 111(37):13475–13480, sep 2014.
- [33] D. D. Ho, A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, 1995.
- [34] Y. C. Ho, L. Shan, N. N. Hosmane, J. Wang, S. B. Laskey, D. I. S. Rosenbloom, J. Lai, J. N. Blankson, J. D. Siliciano, and R. F. Siliciano. XReplication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*, 155(3):540–551, 2013.
- [35] R. D. Hockett, J. Michael Kilby, C. A. Derdeyn, M. S. Saag, M. Sillers, K. Squires, S. Chiz, M. A. Nowak, G. M. Shaw, and R. P. Bucy. Constant Mean Viral Copy Number per Infected Cell in Tissues Regardless of High, Low, or Undetectable Plasma HIV RNA. *The Journal of Experimental Medicine*, 2002.
- [36] B. A. Horsburgh, E. Lee, B. Hiener, J.-S. Eden, T. E. Schlub, S. von Stockenstrom, L. Odevall, J. M. Milush, T. Liegler, E. Sinclair, R. Hoh, E. A. Boritz, D. C. Douek, R. Fromentin, N. Chomont, S. G. Deeks, F. M. Hecht, and S. Palmer. High levels of genetically-intact HIV in HLA-DR+ memory T-cells indicates their value for reservoir studies. *AIDS (London, England)*, jan 2020.
- [37] N. N. Hosmane, K. J. Kwon, K. M. Bruner, A. A. Capoferri, S. Beg, D. I. Rosenbloom, B. F. Keele, Y.-C. Ho, J. D. Siliciano, and R. F. Siliciano. Proliferation of latently infected CD4⁺ T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *The Journal of Experimental Medicine*, 214(4):959–972, 2017.
- [38] P. W. Hunt, H. L. Cao, C. Muzoora, I. Ssewanyana, J. Bennett, N. Emenyonu, A. Kembabazi, T. B. Neilands, D. R. Bangsberg, S. G. Deeks, and J. N. Martin. Impact of CD8 + T-cell activation on CD4 + T-cell recovery and mortality in HIV-infected Ugandans initiating antiretroviral therapy. *AIDS*, 25(17):2123–2131, nov 2011.
- [39] J. Hurst, M. Hoffmann, M. Pace, J. P. Williams, J. Thornhill, E. Hamlyn, J. Meyerowitz, C. Willberg, K. K. Koelsch, N. Robinson, H. Brown, M. Fisher, S. Kinloch, D. A. Cooper, M. Schechter, G. Tambussi, S. Fidler, A. Babiker, J. Weber, A. D. Kelleher, R. E. Phillips, and J. Frater. Immunological biomarkers predict HIV-1 viral rebound after treatment interruption. *Nature Communications*, 6, oct 2015.
- [40] L. Josefsson, S. von Stockenstrom, N. R. Faria, E. Sinclair, P. Bacchetti, M. Killian, L. Epling, A. Tan, T. Ho, P. Lemey, W. Shao, P. W. Hunt, M. Somsouk, W. Wylie, D. C. Douek, L. Loeb, J. Custer, R. Hoh, L. Poole, S. G. Deeks, F. Hecht, and S. Palmer. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proceedings of the National Academy of Sciences of the United States of America*, 110(51):E4987–96, dec 2013.
- [41] J. Z. Li, B. Etemad, H. Ahmed, E. Aga, R. J. Bosch, J. W. Mellors, D. R. Kuritzkes, M. M. Lederman, M. Para, and R. T. Gandhi. The size of the expressed HIV reservoir predicts timing of viral rebound after treatment interruption. *Aids*, 30(3):343–353, jan 2016.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

- [42] P. Lindau, R. Mukherjee, M. V. Gutschow, M. Vignali, E. H. Warren, S. R. Riddell, K. W. Makar, C. J. Turtle, and H. S. Robins. Cytomegalovirus Exposure in the Elderly Does Not Reduce CD8 T Cell Repertoire Diversity. *Journal of immunology (Baltimore, Md. : 1950)*, 202(2):476–483, 2019.
- [43] J. M. Mabuka, A. S. Dugast, D. M. Muema, T. Reddy, Y. Ramlakhan, Z. Euler, N. Ismail, A. Moodley, K. L. Dong, L. Morris, B. D. Walker, G. Alter, and T. Ndung’u. Plasma CXCL13 but not B cell frequencies in acute HIV infection predicts emergence of cross-neutralizing antibodies. *Frontiers in Immunology*, 8(SEP), 2017.
- [44] C. Mackall, T. Fleisher, M. R. Brown, M. M. Andrich, C. C. Chen, I. M. Feuerstein, I. T. Magrath, L. H. Wexler, D. S. Dimitrov, and R. E. Gress. Distinctions between CD8+ and CD4+ T-cell regenerative pathways result in prolonged T-cell subset imbalance after intensive chemotherapy. *Blood*, 89:3700–3707, 1997.
- [45] F. Maldarelli, X. Wu, L. Su, F. R. Simonetti, W. Shao, S. Hill, J. Spindler, A. L. Ferris, J. W. Mellors, M. F. Kearney, J. M. Coffin, and S. H. Hughes. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science*, 345(6193):179–183, 2014.
- [46] M. Markowitz, M. Louie, A. Hurley, E. Sun, M. Di Mascio, A. S. Perelson, and D. D. Ho. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. *Journal of virology*, 77(8):5037–8, apr 2003.
- [47] L. Matrajt, P. M. Younan, H.-P. Kiem, and J. T. Schiffer. The Majority of CD4+ T-Cell Depletion during Acute Simian-Human Immunodeficiency Virus SHIV89.6P Infection Occurs in Uninfected Cells. *Journal of Virology*, 2014.
- [48] W. R. McManus, M. J. Bale, J. Spindler, A. Wiegand, A. Musick, S. C. Patro, M. D. Sobolewski, V. K. Musick, E. M. Anderson, J. C. Cyktor, E. K. Halvas, W. Shao, D. Wells, X. Wu, B. F. Keele, J. M. Milush, R. Hoh, J. W. Mellors, S. H. Hughes, S. G. Deeks, J. M. Coffin, and M. F. Kearney. HIV-1 in lymph nodes is maintained by cellular proliferation during antiretroviral therapy. *The Journal of clinical investigation*, 130:4629–4642, jul 2019.
- [49] R. Mehr and A. S. Perelson. Blind T-cell homeostasis and the CD4/CD8 ratio in the thymus and peripheral blood. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, 14(5):387–398, 1997.
- [50] H. Mohri, A. S. Perelson, K. Tung, R. M. Ribeiro, B. Ramratnam, M. Markowitz, R. Kost, A. Hurley, L. Weinberger, D. Cesar, M. K. Hellerstein, and D. D. Ho. Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *The Journal of experimental medicine*, 194(9):1277–87, nov 2001.
- [51] A. Musick, J. Spindler, E. Boritz, L. Pérez, D. Crespo-Vélez, S. C. Patro, M. D. Sobolewski, M. J. Bale, C. Reid, B. F. Keele, A. Capoferri, W. Shao, A. Wiegand, F. R. Simonetti, J. W. Mellors, S. H. Hughes, J. M. Coffin, F. Maldarelli, and M. F. Kearney. HIV Infected T Cells Can Proliferate in vivo Without Inducing Expression of the Integrated Provirus. *Frontiers in microbiology*, 10:2204, 2019.
- [52] Z. M. Ndhlovu, P. Kamya, N. Mewalal, H. N. Kløverpris, T. Nkosi, K. Pretorius, F. Laher, F. Ogunshola, D. Chopera, K. Shekhar, M. Ghebremichael, N. Ismail, A. Moodley, A. Malik, A. Leslie, P. J. Goulder, S. Buus, A. Chakraborty, K. Dong, T. Ndung’u, and B. D. Walker. Magnitude and Kinetics of CD8+ T Cell Activation during Hyperacute HIV Infection Impact Viral Set Point. *Immunity*, 2015.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

- [53] Z. M. Ndhlovu, S. W. Kazer, T. Nkosi, F. Ogunshola, D. M. Muema, G. Anmole, S. A. Swann, A. Moodley, K. Dong, T. Reddy, M. A. Brockman, A. K. Shalek, T. Ndung'u, and B. D. Walker. Augmentation of HIV-specific T cell function by immediate treatment of hyperacute HIV-1 infection. *Science translational medicine*, 11(493), may 2019.
- [54] D. Nickle. Evolutionary indicators of human immunodeficiency virus type 1 reservoirs and compartments. *Journal of Virology*, 77(5540-5546), 2003.
- [55] T. Oakes, J. M. Heather, K. Best, R. Byng-Maddick, C. Husovsky, M. Ismail, K. Joshi, G. Maxwell, M. Noursadeghi, N. Riddell, T. Ruehl, C. T. Turner, I. Uddin, and B. Chain. Quantitative Characterization of the T Cell Receptor Repertoire of Naïve and Memory Subsets Using an Integrated Experimental and Computational Pipeline Which Is Robust, Economical, and Versatile. *Frontiers in immunology*, 8:1267, 2017.
- [56] A. S. Perelson, D. E. Kirschner, and R. De Boer. Dynamics of HIV infection of CD4 + T cells. *Mathematical Biosciences*, 1993.
- [57] D. Persaud, T. Pierson, C. Ruff, D. Finzi, K. R. Chadwick, J. B. Margolick, A. Ruff, N. Hutton, S. Ray, and R. F. Siliciano. A stable latent reservoir for HIV-1 in resting CD4(+) T lymphocytes in infected children. *The Journal of clinical investigation*, 105(7):995–1003, apr 2000.
- [58] B. Ramratnam, S. Bonhoeffer, J. Binley, A. Hurley, L. Zhang, J. E. Mittler, M. Markowitz, J. P. Moore, A. S. Perelson, and D. D. Ho. Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet*, 1999.
- [59] D. B. Reeves, E. R. Duke, T. A. Wagner, S. E. Palmer, A. M. Spivak, and J. T. Schiffer. A majority of HIV persistence during antiretroviral therapy is due to infected cell proliferation. *Nature communications*, 9(1):4811, 2018.
- [60] D. B. Reeves, C. W. Peterson, H.-P. Kiem, and J. T. Schiffer. Autologous Stem Cell Transplantation Disrupts Adaptive Immune Responses during Rebound Simian/Human Immunodeficiency Virus Viremia. *Journal of Virology*, 91(13), jul 2017.
- [61] R. R. Regoes, D. Wodarz, and M. A. Nowak. Virus dynamics: the effect of target cell limitation and immune responses on virus evolution. *Journal of theoretical biology*, 191(4):451–62, apr 1998.
- [62] R. M. Ribeiro and A. S. Perelson. The Analysis of HIV Dynamics Using Mathematical Models. In *AIDS and Other Manifestations of HIV Infection*. 2004.
- [63] M. L. Robb, L. A. Eller, H. Kibuuka, K. Rono, L. Maganga, S. Nitayaphan, E. Kroon, F. K. Sawe, S. Sinei, S. Sriplienchan, L. L. Jagodzinski, J. Malia, M. Manak, M. S. de Souza, S. To-vanabutra, E. Sanders-Buell, M. Rolland, J. Dorsey-Spitz, M. A. Eller, M. Milazzo, Q. Li, A. Lewandowski, H. Wu, E. Swann, R. J. O'Connell, S. Peel, P. Dawson, J. H. Kim, and N. L. Michael. Prospective Study of Acute HIV-1 Infection in Adults in East Africa and Thailand. *New England Journal of Medicine*, 374(22):2120–2130, 2016.
- [64] B. Roch, N. Dautigny, and P. Pereira. Peripheral T lymphocytes: expansion and potential homeostatic regulation of pool sizes and CD4/CD8 ratios in vivo. *European Journal of Immunology*, 19:905–911, 1989.
- [65] R. Sanjuan and P. Domingo-Calap. Mechanisms of viral mutation. *Cellular and Molecular Life Sciences*, 73:4433–4448, 2016.
- [66] A. R. Schröder, P. Shinn, H. Chen, C. Berry, J. R. Ecker, and F. Bushman. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, 110(4):521–529, aug 2002.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

- [67] S. Sengupta and R. F. Siliciano. Targeting the Latent Reservoir for HIV-1, may 2018.
- [68] L. F. Shampine and M. W. Reichelt. The matlab ode suite. *SIAM Journal on Scientific Computing*, 18(1):1–22, 1997.
- [69] M. B. Shapiro, T. Cheever, D. C. Malherbe, S. Pandey, J. Reed, E. S. Yang, K. Wang, A. Pegu, X. Chen, D. Siess, D. Burke, H. Henderson, R. Lewinsohn, M. Fischer, J. J. Stanton, M. K. Axthelm, C. Kahl, B. Park, A. D. Lewis, J. B. Sacha, J. R. Mascola, A. J. Hessel, and N. L. Haigwood. Single-dose bNAb cocktail or abbreviated ART post-exposure regimens achieve tight SHIV control without adaptive immunity. *Nature communications*, 11(1):70, jan 2020.
- [70] J. D. Siliciano, J. Kajdas, D. Finzi, T. C. Quinn, K. Chadwick, J. B. Margolick, C. Kovacs, S. J. Gange, and R. F. Siliciano. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4 + T cells. *Nature Medicine*, 9(6):727–728, 2003.
- [71] N. Soriano-Sarabia, R. E. Bateson, N. P. Dahl, A. M. Crooks, J. D. Kuruc, D. M. Margolis, and N. M. Archin. Quantitation of Replication-Competent HIV-1 in Populations of Resting CD4+ T Cells. *Journal of Virology*, 88(24):14070–14077, 2014.
- [72] S. Von Stockenstrom, L. Odevall, E. Lee, E. Sinclair, P. Bacchetti, M. Killian, L. Epling, W. Shao, R. Hoh, T. Ho, N. R. Faria, P. Lemey, J. Albert, P. Hunt, L. Loeb, C. Pilcher, L. Poole, H. Hatano, M. Somsouk, D. Douek, E. Boritz, S. G. Deeks, F. M. Hecht, and S. Palmer. Longitudinal Genetic Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool during Effective HIV Therapy. *Journal of Infectious Diseases*, 212(4):596–607, aug 2015.
- [73] T. A. Wagner, J. L. McKernan, N. H. Tobin, K. A. Tapia, J. I. Mullins, and L. M. Frenkel. An Increasing Proportion of Monotypic HIV-1 DNA Sequences during Antiretroviral Treatment Suggests Proliferation of HIV-Infected Cells. *Journal of Virology*, 87(3):1770–1778, feb 2013.
- [74] Z. Wang, E. E. Gurule, T. P. Brennan, J. M. Gerold, K. J. Kwon, N. N. Hosmane, M. R. Kumar, S. A. Beg, A. A. Capoferri, S. C. Ray, Y.-C. Ho, A. L. Hill, J. D. Siliciano, and R. F. Siliciano. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proceedings of the National Academy of Sciences*, 115(11):E2575–E2584, 2018.
- [75] X. Wei, S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*, 1995.
- [76] J. B. Whitney, A. L. Hill, S. Sanisetty, P. Penaloza-Macmaster, J. Liu, M. Shetty, L. Parenteau, C. Cabral, J. Shields, S. Blackmore, J. Y. Smith, A. L. Brinkman, L. E. Peter, S. I. Mathew, K. M. Smith, E. N. Borducchi, D. I. Rosenbloom, M. G. Lewis, J. Hattersley, B. Li, J. Hesselgesser, R. Geleziunas, M. L. Robb, J. H. Kim, N. L. Michael, and D. H. Barouch. Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature*, 512(1):74–77, 2014.
- [77] J. B. Whitney, S. Y. Lim, C. E. Osuna, J. L. Kublin, E. Chen, G. Yoon, P. T. Liu, P. Abbink, E. N. Borducchi, A. Hill, M. G. Lewis, R. Geleziunas, M. L. Robb, N. L. Michael, and D. H. Barouch. Prevention of SIVmac251 reservoir seeding in rhesus monkeys by early antiretroviral therapy. *Nature Communications*, 2018.
- [78] J. P. Williams, J. Hurst, W. Stöhr, N. Robinson, H. Brown, M. Fisher, S. Kinloch, D. Cooper, M. Schechter, G. Tambussi, S. Fidler, M. Carrington, A. Babiker, J. Weber, K. K. Koelsch, A. D. Kelleher, R. E. Phillips, and J. Frater. HIV-1 DNA predicts disease progression and post-treatment virological control. *eLife*, 3:e03821, 2014.

F A. T. Boshier, D. B. Reeves, E. R. Duke, D. A. Swan, M. Prlic, E. F. Cardozo-Ojeda and J. T. Schiffer

- [79] F. Zanini, J. Brodin, L. Thebo, C. Lanz, G. Bratt, J. Albert, and R. A. Neher. Population genomics of inpatient HIV-1 evolution. *eLife*, 4, dec 2015.