Cellular activation, differentiation and proliferation influence the dynamics of genetically-intact proviruses over time

Bethany A. Horsburgh¹,², Bonnie Hiener¹,², Katie Fisher¹,², Eunok Lee¹,², Hannah Morgan³, John-Sebastian Eden¹,², Susanne von Stockenstrom⁴, Lina Odevall⁴, Jeffrey M. Milush⁵, Rebecca Hoh⁵, Rémi Fromentin⁶, Nicolas Chomont⁶, Frederick M. Hecht⁵, Timothy E. Schlub³, Steven G. Deeks⁵, and Sarah Palmer¹,²

¹. Centre for Virus Research, The Westmead Institute of Medical Research, The University of Sydney, Sydney, New South Wales, 2145, Australia
². Sydney Medical School, Westmead Clinical School, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, 2006, Australia
³. Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, 2006, Australia
⁴. Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, 171 77, Sweden
⁵. Department of Medicine, University of California San Francisco, California, 94110, United States of America
⁶. Centre de Recherche du CHUM and Department of Microbiology, Infectiology and Immunology, Université de Montréal, Montreal, H2X 0A9 Canada
**Brief Summary:** The longitudinal analysis of full-length proviruses within memory CD4+ T-cell subsets demonstrates that the biological processes of activation, differentiation and proliferation influence the dynamics of the HIV reservoir, and must be considered during the development of any immune intervention.
Abstract

HIV persists in cells despite antiretroviral therapy, however the influence of cellular mechanisms such as activation, differentiation and proliferation upon the distribution of proviruses over time is unclear. To address this, we used full-length sequencing to examine proviruses within memory CD4+ T-cell subsets longitudinally in eight participants. Over time, the odds of identifying a provirus increased in effector and decreased in transitional memory cells. In all subsets, the more activated (HLA-DR-expressing) cells contained a higher frequency of intact provirus, as did more differentiated cells such as the transitional and effector memory subsets. The proportion of genetically-identical proviruses increased over time, indicating that cellular proliferation was maintaining the persistent reservoir, however, the number of genetically-identical proviral clusters in each subset was stable. As such, key biological processes of activation, differentiation and proliferation influence the dynamics of the HIV reservoir and must be considered during the development of any immune intervention.

Keywords: HIV persistence; HIV reservoir; memory T-cells, cellular proliferation, cellular differentiation, HLA-DR, full-length sequencing
Introduction

For people living with HIV, antiretroviral therapy (ART) must be maintained for a lifetime as replication-competent virus can persist within memory CD4+ T-cells [1-4]. Approximately 1-10% of infected cells contain a provirus with all genetic information necessary to reinitiate infection if therapy ceases [5-9]. These genetically-intact, likely replication-competent proviruses represent the barrier to an HIV cure.

The number of cells infected with replication-competent HIV decays slowly during therapy [10-19]. However, the reservoir is also maintained by cellular proliferation [20]. Antigen-stimulated proliferation can occur when a cell encounters its cognate antigen [21, 22]. By contrast, homeostatic proliferation occurs in response to cytokines, such as IL-7 and IL-15 [23]. Finally, the integrated HIV provirus has also been postulated to induce proliferation [24, 25].

It has been shown that genetically-intact and/or replication-competent HIV is enriched in effector memory (EM) cells [7], Th1 cells [8] and cells displaying activation and exhaustion markers [9, 26-29]. Some of these subsets may be more likely to undergo proliferation, as well as other key processes such as activation or differentiation, which may impact the maintenance of the reservoir during ART [9, 30, 31]. How these cellular processes impact the dynamics of the reservoir within cell subsets over time has not been well-described.

As such, to define the stability of genetically-intact proviruses within naïve and memory T-cell subsets we conducted a longitudinal analysis of full-length proviruses from participants on long-term ART. At the second time-point we co-sorted memory subsets with the HLA-DR
receptor to understand the contribution of activated cells to the reservoir. We also examined the role of cellular proliferation in maintaining persistent HIV over time.

Methods

Support from research institutes

This study was approved by the institutional review board at the Western Sydney Health Department for the Westmead Institute for Medical Research, and the ethics review committees at the University of California San Francisco and Vaccine & Gene Therapy Institute-Florida. Written informed consent was obtained from all participants.

Participants

Leukapheresis samples were obtained from eight participants at two different time-points during ART (Table 1). Time-point one data from four participants was included in previous publications [7, 9], however all data was reanalysed as additional participants were included in this study.

Time-point one: cell sort

Fluorescence-activated cell sorting (FACSaria [BD Biosciences]) was used to sort naïve, Central (CM), Transitional (TM) and Effector Memory (EM) CD4+ T-cell subsets as previously described [7, 12, 30] (Figure S1A). In four participants, CD4+CD45RA-HLA-DR+ and CD4+CD45RA-HLA-DR- memory T-cells were sorted as previously reported [9] (Figure S1B).
Time-point two: cell sort

To obtain naïve, CM±HLA-DR, TM±HLA-DR and EM±HLA-DR cells, peripheral blood mononuclear cells were obtained by Lymphoprep (STEMCELL Technologies) density gradient centrifugation. Total CD4+ T-cells were isolated by magnetic negative selection (STEMCELL Technologies). Cells were then sorted using a FACS Aria [BD Biosciences] using the following antibodies: CD3-FITC (clone UCHT1, BioLegend), CD4-BV650 (clone RPA-T4, BioLegend), CD45RA-PE (clone HI100, BioLegend), CCR7-PECy7 (clone G043H7, BioLegend), CD27-APC (clone M-T271, BioLegend), HLA-DR-BV421 (clone L243, BioLegend) and Live/Dead Aqua Marker (Fisher Bioscience, #L34966) (Figure S1C).

Full-Length Individual Proviral Sequencing (FLIPS)

The FLIPS assay was performed and proviral sequences analysed for defects using a process of elimination [7]. Proviruses not containing a large (>100bp) deletion, inversion, deleterious stop codon, frameshift mutation or mutation in the cis-acting region were classified as genetically-intact.

Identical sequence analysis

Identical sequences were identified using ElimDupes (Los Alamos HIV Database). A cluster of identical sequences was defined as ≥2 100% genetically-identical sequences from the same cell subset.

Statistical methodology

Logistic regression was used to analyse the relationship between cell subset and HIV infection frequency. The contribution of each subset to the reservoir was described using a Wilcoxon signed rank test. The proportion of genetically-identical proviruses within cell
subsets was estimated using a mixed logistic model. Detailed statistical methods can be found in the supplementary data.

Results

The proviral landscape differs between cell subsets

To observe the dynamics of the proviral landscape over time we obtained leukapheresis samples from eight participants at two time-points during ART (Table 1). At time-point one, naïve, CM, TM and EM cells were sorted from seven participants, and HLA-DR± memory CD4+ T-cells sorted from four participants. At time-point two, a median of 4.1 years later (IQR 3.6-4.8), naïve, CM±HLA-DR, TM±HLA-DR and EM±HLA-DR CD4+ T-cells were sorted from all eight participants. We used the FLIPS assay [7] to obtain near full-length (~92%) HIV genomes at the single genome level.

At the first time-point, 1124 genomes were isolated and 48 (4.3%) were genetically-intact (Table S1, Figure S2A). At the second time-point, 1654 genomes were isolated and 105 (6.3%) were genetically-intact (Table S2, Figure S2B). The mutational profile of proviruses differed by subset, with cells considered more differentiated (TM and EM) and cells expressing HLA-DR more likely to contain full-length proviruses, including intact proviruses (Figure S2).

HIV proviruses increase in EM cells and decrease in TM cells with time

To examine the HIV reservoir in different cell subsets we first estimated the overall HIV infection frequency within each subset during ART. At time-point one, there was a significant difference in the infection frequency between subsets, with strong evidence that the nature of this difference was participant dependent (logistic regression model, p<0.0001
for effect modification, Figure 1A). TM cells had the highest estimated infection frequency, 142 cells per million (95% CI 119-168) (Figure 1A, Table S3). All other subsets had an estimated infection frequency of <100 cells per million, with a progression TM>CM>EM>naïve (Figure 1A, Table S3).

At the second time-point, the total infection frequency was again significantly different between subsets, with strong evidence for participant effect modification (logistic regression model, p<0.0001, Figure 1B). EM/HLA-DR+ cells had the highest estimated infection frequency, 203 cells per million (95% CI 165-251) and naïve cells had the lowest, 14 infected cells per million (95% CI 13-16) (Figure 1B, Table S3). All other subsets had an estimated infection frequency between 58-75 cells per million (Figure 1B, Table S3).

To assess if the infection frequencies within each subset changed over four years, the CM, TM and EM subsets at time-point two were recapitulated using the biological proportions of HLA-DR+ and HLA-DR- cells in each subset (Table S4). A 2-fold decrease in the infection frequency of TM cells was observed (odds ratio=0.53; p<0.0001) (Figure 1C, Table S5). However, HIV-infected EM cells increased by 28% (odd ratio=1.28; p=0.002) (Figure 1C, Table S5). There was no significant change in the frequency of infected naïve or CM cells with time (Figure 1C, Table S5)

Genetically-intact HIV proviruses are concentrated in differentiated cells and/or cells expressing HLA-DR

To understand the dynamics of intact proviruses, we then assessed the change in the infection frequency of intact provirus within cell subsets over time. At time-point one there was a significant difference in the infection frequency with intact provirus between subsets
(p<0.0001) (Figure 2A). The estimated frequency of intact proviruses ranged from 0.2-3 per million cells, and increased with differentiation status (CM<TM<EM) (Figure 2A, Table S6). At the second time-point there was also a significant difference (p<0.0001) in the infection frequency between the subsets (Figure 2B); in all subsets an estimated 0-3 cells per million contained an intact provirus, except for EM/HLA-DR+ cells which had an estimated infection frequency of 31 cells per million (95% CI 12-84) (Figure 2B, Table S6). However, only two participants had high levels of intact provirus in their EM/HLA-DR+ cells. A stepwise progression was observed in terms of the frequency of intact provirus within the subsets; for both HLA-DR- and HLA-DR+ cells this infection frequency progressed CM<TM<EM (Figure 2B). Moreover, in each subset HLA-DR- cells had a lower estimated frequency of intact proviruses than HLA-DR+ cells (Figure 2B, Table S6).

We then compared the genetically-intact infection frequencies for each subset between time-points. After recapitulating the time-point two memory subsets from the HLA-DR+/−components (Table S4), we found that the intact infection frequency did not change significantly in any subset over time (Figure 2C; Table S7).

Memory CD4+ T-cell subsets do not contribute equally to the peripheral blood HIV reservoir

As the proportion of each cell subset within the peripheral blood is different, we assessed the contribution of each subset to the pool of infected cells during therapy taking these proportions into account. That is, we used the biological proportions of each subset within the pool of peripheral blood CD4+ T-cells, and the infection frequency within that subset, to calculate their contribution to the reservoir. We assessed this at time-point two only as each participant had the same subsets sorted, and the contribution of time-point one cell subsets was determined in previous studies [7, 9].
Over half (median 59.9%, IQR 45.8-55.3%) of all HIV proviruses, defective or intact, were found in the CM/HLA-DR- cell subset (Figure 3A), significantly higher than any other subset (Table S8). Naïve cells contributed a further 11.5% (IQR 5.9-15.6%) (Figure 3A), however this contribution was only significantly higher than CM/HLA-DR+ cells (p=0.04) (Table S8). Naïve and CM/HLA-DR- cells were the two most prevalent peripheral blood CD4+ T-cell subsets (Table S9), accounting for their high contribution. All other subsets contributed <10% of the total HIV burden at this time. However, we observed a large amount of variability in this contribution across participants (Figure 3A).

We then compared the relative contribution of the cell subsets to the genetically-intact proviral reservoir in the peripheral blood at time-point two. The highest contributor was found to be EM/HLA-DR- cells, contributing a median of 39.2% (IQR 4.3-43.3). Despite their low infection frequency, naïve cells were also found to be a high contributor, as has been observed previously [19], contributing a median of 13.7% (IQR 17.2-1.4) (Figure 3B). All other subsets contributed ≤10% to the genetically-intact reservoir. However, no subset was found to contribute significantly more than any other (Table S10), likely due to the small numbers of intact proviruses found.

Finally, as we had previously observed that HLA-DR+ memory CD4+ T-cells contributed over half of all intact proviruses to the peripheral blood reservoir [9], we wanted to establish how these cells contributed to the reservoir found within specific memory cell subsets. We observed that the contribution of HLA-DR+ cells to the total pool of infected cells increased as cells differentiated (Figure 3C). Significantly more CM/HLA-DR- cells contributed to the overall reservoir compared to CM/HLA-DR+ cells (p=0.008, Wilcoxon signed rank test), however in the EM subset HLA-DR+ and HLA-DR- cells contributed approximately equally.
(Figure 3C). When we assessed the contribution of these cells to the genetically-intact reservoir within each subset, we found that HLA-DR+ cells were the highest contributor to the CM subset (Figure 3D), though not significantly. Interestingly, these genetically-intact proviruses were not unique to this subset, as 5/7 of the intact proviruses from CM/HLA-DR+ cells were also found within EM cells in the same participant (Figure 3D). By contrast, the contribution of HLA-DR+ cells was lower than HLA-DR- cells in the TM and EM subsets, though not significantly (Figure 3D).

**Cellular proliferation maintains HIV proviruses within the peripheral blood reservoir**

We then sought to understand how HIV proviruses were maintained over four years of therapy. As intact proviruses may be seeded by the infection of previously uninfected cells, we looked for evidence of this by calculating the change in genetic diversity of intact proviruses over time. In three participants where >5 genetically-intact proviruses were isolated at both time-points, there was no change in genetic diversity with time (p=0.20) – in fact, genetic diversity decreased in two participants – indicating new infections were unlikely to be occurring (Figure S3).

To determine if proviruses were being maintained by cellular proliferation, we screened for 100% genetically-identical proviruses. We note that while homogenous proviral populations may be formed if ART is initiated during acute/early infection, there was no difference in the proportion of genetically-identical proviruses between participants treated during acute/early or chronic infection (time-point one, p=0.11; time-point two, p=0.15) (Figure S4). As such, the presence of genetically-identical proviruses within this study were unlikely to be influenced by the timing of ART initiation.
Examples of identical genomes were found in all subsets, and all participants had at least one example of a cluster of genetically-identical proviruses. At time-point one, the proportion of identical sequences was different between subsets, with evidence that the nature of this difference was participant-dependent (logistic regression, p<0.0001 for effect modification). Overall, more differentiated subsets were observed to have a higher proportion of genetically-identical proviruses (Figure 4A). This pattern was maintained at time-point two, as we again found that the proportion of identical sequences was significantly different between cell subsets, with evidence that the nature of this difference was participant-dependent (logistic regression, p<0.0001 for effect modification). A stepwise progression was observed as cells differentiated, CM<TM<EM. Moreover, a higher proportion of identical genomes was observed in HLA-DR+ cells compared to HLA-DR- cells of the same subset (Figure 4B).

After taking into account the biological proportions of HLA-DR+ and HLA-DR- cells at time-point two (Table S4), we found that the overall proportion of identical sequences increased over time in all subsets, though not significantly (Table S11).

We then assessed how proliferation may be impacting the landscape of intact proviruses over time. Overall, 23% and 58% of genetically-intact sequences at time-points one and two were identical. All three participants with >5 intact HIV proviruses isolated at both time-points showed an increase in the proportion of genetically-identical intact proviruses over time (p=0.037) (Figure S5). At time-point one, genetically-identical intact proviruses were identified from two participants only, however the proportion of these increased as cells differentiated (Figure 5A). Similarly, at time-point two, a stepwise progression in terms of the proportion of genetically-identical intact proviruses was observed in both the HLA-DR- and HLA-DR+ subsets, EM>TM>CM (Figure 5B). However, we note that interpretations of this data are limited due to the rarity of these sequences.
Interestingly, clusters of identical sequences may have impacted our ability to identify intact proviruses in some participants. No intact provirus was isolated from three participants at either time-point; however all three participants had one dominant expansion of defective sequences within their reservoir that was observed at both time-points and in all memory cell subsets (Figure S6).

The identity of cells undergoing cellular proliferation changes with time

We then investigated the number of unique clusters of identical sequences in each subset - likely representing different proliferative events - to understand the dynamics of identical sequences with time. At both time-points we found that the number of clusters of identical sequences, both overall (Figure 6) and intact (Figure S7), was higher in more differentiated cells, though variability in the number of clusters was observed between participants.

We then compared the number of clusters of identical sequences within each subset over time to determine if an increase of these clusters may have contributed to the increase in the proportion of identical sequences with time. The number of clusters observed in TM cells was different between time-points (p=0.0006), with all participants having an increase of clusters at the second time-point (Figure S8). There was no change in the number of clusters of identical sequences in any other subset (Figure S8). Given this stability, we wanted to understand if this was reflected in the proviral populations that were identical at each time-point. However, we found that the populations isolated at time-points one and two were different, reflecting underlying fluctuations in the makeup of the reservoir (Table S12-S18).
Discussion

In this study, we used full-length sequencing of the HIV genome to analyse the dynamics of the HIV reservoir within CD4+ T-cell subsets over time during therapy. At time-point one, TM and EM cells contained the highest frequencies of genetically-intact provirus. This was not surprising as the importance of differentiated cells in harbouring genetically-intact and/or translation-competent virus has been demonstrated previously [7, 26, 32]. However, this pattern was also observed at a second time-point, suggesting that intact HIV genomes may be stably maintained within differentiated cells with time. As EM cells are known to have a shorter lifespan compared to CM [33, 34], there may be a compensatory process that aids in the maintenance of the reservoir. One such mechanism may be the replenishment of the reservoir through the differentiation of infected naïve, CM or TM cells into EM with time. In particular, the finding of intact proviruses within long-lived naïve cells at both time-points suggests that naïve cell differentiation into TM and EM cells after stimulation may be a source of persistent HIV during long-term therapy [19]. While we observed total proviruses to decrease in TM cells and increase in EM cells between time-points – suggesting that infected cells move between subsets – the low frequency of cells infected with intact provirus meant we were unable to directly observe the differentiation of cells infected with genetically-intact HIV in this study.

Previously we observed that HLA-DR+ memory cells were enriched for genetically-intact provirus [9]. Other studies have shown that markers such as PD-1 [31, 35], TIGIT [28] and OX40 [29] are both expressed on cells that contain a high frequency of provirus and correlate with HLA-DR expression. As such, similar concentrations of HIV within HLA-DR+ memory cells may be likely. However, given the overlap between differentiated cells and HLA-DR+
cells, we were previously unable to determine whether a high proportion of intact genomes in HLA-DR+ cells was due to cellular activation or differentiation [9]. However, within this study there was a higher frequency of genetically-intact provirus in HLA-DR+ cells compared to HLA-DR- in all memory cell subsets, indicating both activation and differentiation play a role in maintaining the HIV reservoir. In particular, cellular activation may be important for maintaining the reservoir in less differentiated cell subsets that do not otherwise contain high frequencies of intact provirus, as demonstrated by the high contribution of HLA-DR+ cells to the intact reservoir in the CM subset.

Cellular proliferation may be a third intrinsic cellular process maintaining the viral reservoir [8, 20, 22-25, 30, 36]. Notably, there may be a strong interplay between proliferation, differentiation and activation, as examples of identical proviruses were increased in differentiated and/or activated cell subsets at both time-points. These identical proviruses were found to increase in all subsets between time-points, though not significantly, in contrast to previous studies [37]. Notably, we also found that intact, genetically-identical proviruses increased over time, further indicating that cellular proliferation is maintaining the latent reservoir. Recent work has highlighted that intact proviruses undergoing proliferation may not be targeted by the immune system for depletion [16] and instead that cellular proliferation maintains both intact and defective proviruses. Moreover, additional studies have shown that the administration of IL-7 also increases the absolute number of infected cells in the reservoir through an increase in cell proliferation [23]. As such, cell proliferation may have the potential to not only maintain, but expand, the reservoir with time.

Finally, similar to what has been observed previously [14, 30, 38], while the number of clusters of identical sequences was stable over time, the identity of these populations
changed. As cytokine-driven proliferation is not selective for specific cell populations [38], shifts in the dominant population may be predominately driven by antigenic proliferation. Recent work has highlighted that cell populations reactive to common antigens harbour proviruses during prolonged therapy [21, 22, 36]. During the course of any immune response, these cells (and therefore integrated proviruses) may expand and contract [22], similar to the fluctuations observed both here and previously [38]. As such, changes in the clusters of proviruses that are observed over time may be a reflection of the larger dynamics of the immune system.

However, while inferences can be drawn regarding the roles of cellular activation, differentiation and proliferation in shaping the reservoir over time, it is important to note that the small number of participants (n=8) limits the extent to which we can observe patterns within the study, with participant effect modification observed in many of our statistical analysis. Moreover, the number of intact proviruses able to be isolated was limited – including three participants with no intact provirus isolated – meaning patterns involving intact proviruses must be interpreted with caution. Furthermore, while there may be a biological basis for the patterns observed herein, evidence of participant effect modification may indicate that there is an additional stochastic nature to this distribution.

Overall, genetically-intact HIV proviral infection frequencies are increased in more differentiated and activated cell subsets during long-term ART. Increased proportions of identical sequences are also observed within the same subsets over time. However, underlying shifts in the proviral landscape are evident through an increase in the proportion of identical sequences with time and the different proviruses contributing to these identical sequence clusters between time-points. Given that the proliferation of each cell subset –
likely driven by the immune response – is influencing the dynamics of the HIV reservoir, clearly taking cellular biology into account – including the cell subset of origin of the provirus and therefore the likelihood of that cell to proliferate with time – is crucial for a complete understanding of the persistent HIV reservoir, and thereby the development of an effective and efficient HIV cure.
References

1. Wong JK, Hezareh M, Gunthard HF, et al. Recovery of Replication-Competent HIV Despite Prolonged Suppression of Plasma Viremia. Science 1997; 278:1291-5.

2. Finzi D, Blankson J, Siliciano JD, 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 1999; 5:512-7.

3. Finzi D, Hermankova M, Pierson T, et al. Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral Therapy. Science 1997; 278:1295-300.

4. Chun T-W, Stuyver L, Mizell SB, et al. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proceedings of the National Academy of Sciences of the United States of America 1997; 94:13193-7.

5. Ho Y-C, Shan L, Hosmane Nina N, et al. Replication-Competent Noninduced Proviruses in the Latent Reservoir Increase Barrier to HIV-1 Cure. Cell 2013; 155:540-51.

6. Bruner KM, Murray AJ, Pollack RA, et al. Defective proviruses rapidly accumulate during acute HIV-1 infection. Nat Med 2016; 22:1043-9.

7. Hiener B, Horsburgh BA, Eden J-S, et al. Identification of Genetically Intact HIV-1 Proviruses in Specific CD4+ T Cells from Effectively Treated Participants. Cell Reports 2017; 21:813-22.

8. Lee GQ, Orlova-Fink N, Einkauf K, et al. Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. The Journal of Clinical Investigation 2017; 127:2689-96.

9. Horsburgh BA, Lee E, Hiener B, et al. High levels of genetically intact HIV in HLA-DR+ memory T cells indicates their value for reservoir studies. AIDS 2020; 34:659-68.
10. Siliciano JD, Kajdas J, Finzi D, et al. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nature Medicine 2003; 9:727+.

11. Palmer S, Maldarelli F, Wiegand A, et al. Low-level viremia persists for at least 7 years in patients on suppressive antiretroviral therapy. Proceedings of the National Academy of Sciences of the United States of America 2008; 105:3879-84.

12. Josefsson L, von Stockenstrom S, Faria NR, et al. 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 2013; 110:E4987-E96.

13. Besson GJ, Lalama CM, Bosch RJ, et al. HIV-1 DNA Decay Dynamics in Blood During More Than a Decade of Suppressive Antiretroviral Therapy. Clinical Infectious Diseases 2014; 59:1312-21.

14. Pinzone MR, VanBelzen DJ, Weissman S, et al. Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by clonal expansion. Nature Communications 2019; 10:728.

15. Bruner KM, Wang Z, Simonetti FR, et al. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. Nature 2019; 566:120-5.

16. Antar AA, Jenike KM, Jang S, et al. Longitudinal study reveals HIV-1-infected CD4+ T cell dynamics during long-term antiretroviral therapy. J Clin Invest 2020; 130:3543-59.

17. Gandhi RT, Cyktor JC, Bosch RJ, et al. Selective Decay of Intact HIV-1 Proviral DNA on Antiretroviral Therapy. The Journal of Infectious Diseases 2020.

18. Peluso MJ, Bacchetti P, Ritter KD, et al. Differential decay of intact and defective proviral DNA in HIV-1–infected individuals on suppressive antiretroviral therapy. JCI Insight 2020; 5.
19. Venanzi Rullo E, Pinzone MR, Cannon L, et al. Persistence of an intact HIV reservoir in phenotypically naive T cells. JCI Insight 2020; 5.

20. Chomont N, Far ME, Ancuta P, et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. Nature Medicine 2009; 15:893+.

21. Mendoza P, Jackson JR, Oliveira TY, et al. Antigen-responsive CD4+ T cell clones contribute to the HIV-1 latent reservoir. J Exp Med 2020; 217.

22. Gantner P, Pagliuzza A, Pardons M, et al. Single-cell TCR sequencing reveals phenotypically diverse clonally expanded cells harboring inducible HIV proviruses during ART. Nature communications 2020; 11:4089-.

23. Vandergeeten C, Fromentin R, DaFonseca S, et al. Interleukin-7 promotes HIV persistence during antiretroviral therapy. Blood 2013; 121:4321-9.

24. Wagner TA, McLaughlin S, Garg K, et al. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. Science 2014; 345:570-3.

25. Maldarelli F, Wu X, Su L, et al. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. Science 2014; 345:179-83.

26. Pardons M, Baxter AE, Massanella M, et al. Single-cell characterization and quantification of translation-competent viral reservoirs in treated and untreated HIV infection. PLOS Pathogens 2019; 15:e1007619.

27. Fromentin R, Bakeman W, Lawani MB, et al. CD4+ T Cells Expressing PD-1, TIGIT and LAG-3 Contribute to HIV Persistence during ART. PLOS Pathogens 2016; 12:e1005761.

28. Chew GM, Fujita T, Webb GM, et al. TIGIT Marks Exhausted T Cells, Correlates with Disease Progression, and Serves as a Target for Immune Restoration in HIV and SIV Infection. PLOS Pathogens 2016; 12:e1005349.

29. Kuo H-H, Ahmad R, Lee GQ, et al. Anti-apoptotic Protein BIRC5 Maintains Survival of HIV-1-Infected CD4+ T Cells. Immunity 2018; 48:1183-94.e5.
30. von Stockenstrom S, Odevall L, Lee E, et al. Longitudinal Genetic Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool During Effective HIV Therapy. Journal of Infectious Diseases 2015; 212:596-607.

31. Lee E, Bacchetti P, Milush J, et al. Memory CD4 + T-Cells Expressing HLA-DR Contribute to HIV Persistence During Prolonged Antiretroviral Therapy. Frontiers in Microbiology 2019; 10.

32. Neidleman J, Luo X, Frouard J, et al. Phenotypic analysis of the unstimulated in vivo HIV CD4 T cell reservoir. eLife 2020; 9:e60933.

33. Macallan DC, Wallace D, Zhang Y, et al. Rapid turnover of effector-memory CD4(+) T cells in healthy humans. The Journal of experimental medicine 2004; 200:255-60.

34. Bacchus-Souffan C, Fitch M, Symons J, et al. Relationship between CD4 T cell turnover, cellular differentiation and HIV persistence during ART. PLOS Pathogens 2021; 17:e1009214.

35. Psomas C, Younas M, Reynes C, et al. One of the immune activation profiles observed in HIV-1-infected adults with suppressed viremia is linked to metabolic syndrome: The ACTIVIH study. EBioMedicine 2016; 8:265-76.

36. Simonetti FR, Zhang H, Soroosh GP, et al. Antigen-driven clonal selection shapes the persistence of HIV-1–infected CD4+ T cells in vivo. The Journal of Clinical Investigation 2021; 131.

37. Cohn LB, Silva IT, Oliveira TY, et al. HIV-1 integration landscape during latent and active infection. Cell 2015; 160:420-32.

38. Wang Z, Gurule EE, Brennan TP, et al. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. Proceedings of the National Academy of Sciences of the United States of America 2018; 115:E2575-E84.
Table Legend

Table 1

* For four participants where indicated, the time to viral suppression (<40 copies/mL plasma) after ART administration was unknown at time-point one, so for time-point one the time listed is years on ART

3TC, lamivudine; ABC, abacavir; AZT, zidovudine; COBI, cobicistat; DRV, darunavir; ECV, entecavir; EFV, efavirenz; EGV, elvitegravir; ETR, Etravirine; FPV, fosamprenavir; FTC, emtricitabine; MVC, maraviroc; NVP, nevirapine; RPV, rilpivirine; RTG, raltegravir; RTV, ritonavir; TAF, tenofovir alafenamide; TCV, dolutegravir; TDF, tenofovir disoproxil fumarate

CM, Central Memory; TM; Transitional Memory; EM, Effector Memory
Figure Legends

Figure 1: Infection frequencies of cell subsets with any HIV provirus, intact or defective, in the peripheral blood. (A). Infection frequency of cell subsets at time-point one. Logistic regression model, estimated total infection frequency ± 95% CI. (B). Infection frequency of cell subsets at time-point two. Logistic regression model, estimated mean ± 95% CI. (C). Change in the infection frequency of cell subsets between time-points one and two. Memory cell subsets for time-point two recapitulated from the biological proportions of HLA-DR± cells in each subset. Pairwise comparisons, estimated mean ± 95% CI.
CM: Central Memory TM: Transitional Memory EM: Effector Memory

Figure 2: Infection frequencies of cell subsets with an intact HIV provirus in the peripheral blood. (A). Infection frequency with an intact provirus within differentiated cell subsets at time-point one. Logistic regression model, estimated intact infection frequency ± 95% CI. (B). Infection frequency with an intact provirus in differentiated cell subsets at time-point two. Logistic regression model, estimated mean ± 95% CI. (C). Change in the infection frequency with an intact provirus within differentiated cell subsets between time-points one and two. Cell subsets for time-point two recapitulated from biological proportions. Pairwise comparisons, estimated mean ± 95% CI. CM: Central Memory TM: Transitional Memory EM: Effector Memory

Figure 3. Contribution of memory cell subsets to the burden of HIV in the peripheral blood. (A). Contribution of cell subsets to the population of CD4+ T-cells containing any HIV provirus in the peripheral blood at time-point two. Data is median ± interquartile range. (B). Contribution of cell subsets to the intact HIV reservoir in the peripheral blood at time-
point two. Data is median ± interquartile range. (C). Contribution of cells that are HLA-DR+ or HLA-DR- to the total population of each memory cell phenotype that is infected with any HIV provirus at time-point two. Data is median ± interquartile range. Wilcoxon signed rank test. (D). Contribution of cells that are HLA-DR+ or HLA-DR- to the population of each memory cell phenotype that contains a genetically-intact HIV provirus at time-point two. Data is median ± interquartile range. Wilcoxon signed rank test.

Figure 4. Proportion of identical sequences isolated from different cell subsets at time-point one (A) and time-point two (B). Data is mean ± 95% confidence interval.

Figure 5. Proportion of genetically-intact, genetically-identical sequences isolated from different cell subsets at time-point one (A) and time-point two (B). Data is mean ± S.D.

Figure 6. The number of clusters of genetically-identical sequences (intact or defective) isolated from different cell subsets at time-point one (A) and time-point two (B). Data is mean ± S.D.
Conflicts of Interest

NC has received consulting fees from Gilead Sciences, and the laboratory of NC has received a grant from EMD Serono. All other authors have nothing to declare.

Acknowledgements

We would like to thank the Sydney Informatics Hub (University of Sydney, Sydney, Australia) for support for and access to analytical software, and the Ramaciotti Centre for Genomics (University of New South Wales, Sydney, Australia) and the Australian Genome Research Facility (AGRF Sydney Site, the Westmead Institute for Medical Research, Sydney, Australia) for operating the MiSeq. We acknowledge with gratitude the participants who donated samples for this study.

Funding

This work was supported by the Delaney AIDS Research Enterprise (DARE) to Find a Cure (1U19AI096109 and 1UM1AI126611-01), an amfAR Research Consortium on HIV Eradication (ARCHE) Collaborative Research Grant from The Foundation for AIDS Research (amfAR 108074-50-RGRL), the UCSF/Gladstone Institute of Virology & Immunology Center for AIDS Research (P30 AI027763), Australian Centre for HIV and Hepatitis Virology Research (ACH2) (2015-43), and the Australian National Health and Medical Research Council (APP1061681 and APP1149990).

Previous Presentations

Sections of this work have been presented at the Conference on Retroviruses and Opportunistic Infections 2020, and AIDS2020:Virtual.
Corresponding Author Information

Bethany A. Horsburgh

The Westmead Institute of Medical Research, 176 Hawkesbury Road, Westmead, Australia 2145; Phone: +61 2 8627 3000; email: bethany.horsburgh@sydney.edu.au

Alternative Corresponding Author Information

Sarah Palmer

The Westmead Institute of Medical Research, 176 Hawkesbury Road, Westmead, Australia 2145; Phone: +61 2 8627 3620; email: sarah.palmer@sydney.edu.au

Author Contributions

B.A.H., B.H and K.F conducted FLIPS on participant samples and analysed data; R.H., S.D., F.H., J.M.M., S.v.S., L.O., N.C., and R. F. enrolled the participants, collected and/or sorted cell subsets from the participant samples; B.A.H, B.H., K.F., E.L., S.v.S. and L.O. prepared participant samples; J-S.E. designed analysis and data visualization workflows; H.M and T.E.S. conducted the statistical analyses. B.A.H. wrote the original manuscript; S.P. designed the study, supervised the work performed, and edited the manuscript.

Data and Materials Availability

Genbank accession numbers are MH843739-MH843916, KY778264-KY778681, KY766150-KY766212 and MZ080627-MZ081008 (time-point one); MW753232-MW754712 (time-point two).
Table 1. Participant Characteristics

| Participant | Sex | Age (Time before Initiation of Therapy) | Time Before Initiation of Therapy (months) | CD4 Nadir | Years Between Time-Points | Leukapheresis One (Time-Point One) | Leukapheresis Two (Time-Point Two) |
|-------------|-----|----------------------------------------|-------------------------------------------|-----------|--------------------------|-----------------------------------|-----------------------------------|
|             |     |                                        |                                           |           |                          | Treatment Regime                  | Treatment Regime                  |
|             |     |                                        |                                           |           |                          | CD4 + T-Cell Count               | Viral Load (Copies/mL)            | CD4 + T-Cell Count               | Viral Load (Copies/mL)            | Cells Available |
| 2026        | M   | 63                                     | >12                                       | 132       | 3.7                      | TDF, ABC/3TC, RTV, DRV           | 4/18                             | <40                              | Naive, CM, TM, EM, HLA-DR-      | 16.9 | ABC/TCV/3TC | 414 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |
| 2046        | M   | 53                                     | >12                                       | 10        | 2.8                      | ECV, EFV/TDF/FTC                 | 1099                             | <40                              | Naive, CM, TM, EM, HLA-DR-      | 18.5 | EFV/TDF/FTC, ECV | 528 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |
| 1292        | M   | 49                                     | >12                                       | 342       | 4.6                      | EFV/FTC/TDF                      | 746                              | <40                              | Naive, CM, TM, EM               | 7.9  | EFV/TDF/FTC | 677 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |
| 2518        | F   | 59                                     | >12                                       | 70        | 3.5                      | TDF, AZT/3TC, NVP                | 432                              | <40                              | Naive, CM, TM, EM, HLA-DR-      | 18.3 | TDF, AZT/3TC, NVP | 179 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |
| 2013        | M   | 70                                     | >12                                       | 14        | 4.0                      | ABC/3TC, RGV                     | 819                              | <40                              | Naive, CM, TM, EM, HLA-DR-      | 21.2 | ABC/TCV/3TC | 530 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |
| 2303        | M   | 44                                     | <6                                        | 478       | 4.2                      | EFV/FTC/TDF                      | 606                              | <40                              | Naive, CM, TM, EM               | 8.9  | ABC/TCV/3TC | 880 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |
| 2302        | M   | 32                                     | <6                                        | 391       | 4.9                      | FPV, RTV, TDF/FTC               | 696                              | <40                              | Naive, CM, TM, EM               | 9.5  | EGV/TAF/FTC/COBI | 654 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |
| 2286        | M   | 52                                     | <6                                        | 165       | 5.4                      | EFV/FTC/TDF                     | 381                              | <40                              | Naive, CM, TM, EM               | 14.5 | RPV/TDF/FTC | 400 | <40 | Naive, CM, HLA-DR- | EM/HLA-DR- |

* For four participants where indicated, the time to viral suppression (<40 copies/mL plasma) after ART administration was unknown at time-point one, so for time-point one the time listed is years on ART

3TC, lamivudine; ABC, abacavir; AZT, zidovudine; COBI, cobicistat; DRV, darunavir; ECV, entecavir; EFV, efavirenz; EGV, elvitegravir; ETR, Etravirine; FPV, fosamprenavir; FTC, emtricitabine; MVC, maraviroc; NVP, nevirapine; RPV, rilpivirine; RTG, raltegravir; RTV, ritonavir; TAF, tenofovir alafenamide; TCV, dolutegravir; TDF, tenofovir disoproxil fumarate

CM, Central Memory; TM, Transitional Memory; EM, Effector Memory
Figure 2

Intact HIV proviruses per 10^6 cells

A. [Graph showing data points and statistical values.]

B. [Graph with similar data points and statistical values.]

C. [Graph with additional data points and statistical values.]
Figure 3

A. Contribution to overall reservoir [%]

B. Contribution to the intact proviral reservoir [%]

C. Contribution to overall reservoir within subset [%]

D. Contribution to the intact proviral reservoir within subset [%]
Figure 4

A. difference between all subsets

\[ p < 0.001 \]

% Identical sequences

- Naive
- CM
- TM
- EM

B. difference between all subsets

\[ p < 0.0001 \]

% Identical sequences

- Naive
- CM\(\text{HLA-DR-}\)
- CM\(\text{HLA-DR+}\)
- TM\(\text{HLA-DR-}\)
- TM\(\text{HLA-DR+}\)
- EM\(\text{HLA-DR-}\)
- EM\(\text{HLA-DR+}\)
Figure 5

A. Proportion of intact identical sequences (%)

B. Proportion of intact identical sequences (%)

Time Point One

Time Point Two
