

HIV persists throughout deep tissues with repopulation from multiple anatomical sources

Antoine Chaillon, ... , Bram Vrancken, Davey M. Smith

J Clin Invest. 2020. <https://doi.org/10.1172/JCI134815>.

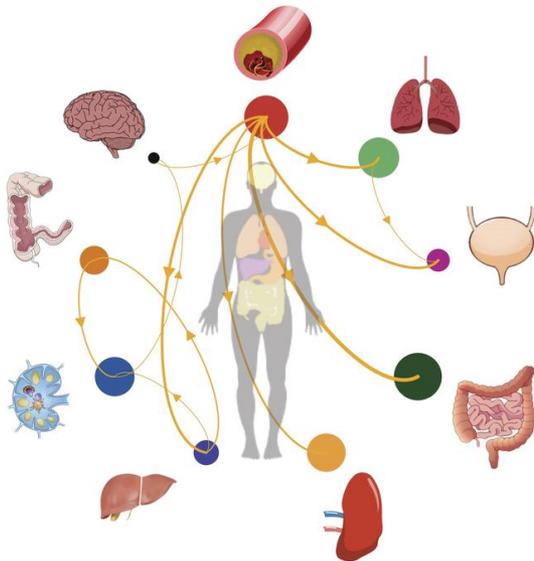
Clinical Research and Public Health

In-Press Preview

AIDS/HIV

Infectious disease

Graphical abstract



Find the latest version:

<https://jci.me/134815/pdf>



HIV persists throughout deep tissues with repopulation from multiple anatomical sources

Antoine Chaillon^{1*}, Sara Gianella^{1*}, Simon Dellicour^{2,3}, Stephen A Rawlings¹, Timothy E. Schlub⁴, Michelli Faria De Oliveira¹, Caroline Ignacio¹, Magali Porrachia¹, Bram Vrancken^{3^} and Davey M Smith^{1^}

1. Department of Medicine, University of California San Diego, La Jolla, California, United States of America.
2. Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, CP160/12 50, av. FD Roosevelt, 1050 Bruxelles, Belgium
3. KU Leuven, Department of Microbiology and Immunology, Rega Institute, Laboratory of Computational and Evolutionary Virology, Leuven, Belgium;
4. University of Sydney, Faculty of Medicine and Health, Sydney School of Public Health, Sydney 2000, NSW, Australia

*contributed equally

^contribute equally

Keywords

Discrete phylogeography; generalized linear model (GLM); viral dynamics; within host; reservoirs.

Corresponding author and requests for reprints:

Antoine Chaillon, M.D. Ph.D.
Department of Medicine
University of California San Diego
9500 Gilman Drive MC 0679
La Jolla, CA 92093-0679, USA
Tel: (858) 552-8585 #2673
Fax: (858) 552-7445
E-mail: achaillon@ucsd.edu

ABSTRACT

Background. Understanding HIV dynamics across the human body is important for cure efforts. This goal has been hampered by technical difficulties and the challenge to obtain fresh tissues.

Methods. This observational study evaluated 6 persons with HIV (4 virally suppressed with antiretroviral therapy and 2 with rebound viremia after stopping therapy) who provided blood serially before death and their bodies for rapid autopsy. HIV reservoirs were characterized by digital droplet PCR and single genome amplification and sequencing of full-length (FL) envelope HIV. Phylogeographic methods reconstructed HIV spread and generalized linear models tested for viral factors associated with dispersal.

Results. Across participants, HIV DNA levels varied from ~0 to 659 copies/ 10^6 cells (IQR:22.9-126.5). A total of 605 intact FL *env* sequences were recovered in *antemortem* blood cells and across 28 tissues (IQR:5-9). Sequence analysis showed: 1) emergence of large, identical, intact HIV RNA populations in blood after stopping therapy, which repopulated tissues throughout the body, 2) multiple sites acted as hubs for HIV dissemination but blood and lymphoid tissues were the main source, and 3) viral exchanges occurred within brain areas and across the blood brain barrier, and 4) migration was associated with low HIV divergence between sites and higher diversity at the recipient site.

Conclusion. HIV reservoirs persist in all deep tissues, and blood is the main source of dispersal. This may explain why eliminating HIV susceptibility in circulating T cells via bone marrow transplants allowed some people with HIV to have therapy free remission, even though deeper tissue reservoirs were not targeted.

Trial registration. Not applicable.

Funding. National Institute of Health Grants (P01 AI31385, P30 AI036214, AI131971-01, AI120009AI036214,HD094646, AI027763, AI134295, AI68636).

INTRODUCTION

Upon transmission to a new host, HIV-1 disseminates to the lymph nodes within days, and to the bloodstream and throughout the human body within 1-2 weeks(1-4). In this process, HIV establishes reservoirs throughout the body including: the central nervous system (CNS), lymphoid tissues (spleen, thymus, lymph nodes, gut associated lymphoid tissue), bone marrow, lungs, kidneys, liver, adipose tissue, gastrointestinal tract and genitourinary systems(5). The persistence of replication competent HIV in these anatomical sites, even in the setting of longstanding, potent antiretroviral therapy (ART) that suppresses HIV replication, is the main barrier to curing HIV(6).

Despite extensive investigations in humans (5, 7-18) much remains unclear about HIV reservoirs that persist during ART(19-23). In part, this is because of technical limitations and limited access to appropriately collected tissues for such studies. In general, most studies aiming to characterize HIV DNA reservoirs in persons with HIV (PWH) have necessarily focused on blood with a few other compartments that are relatively easy to sample, such as the gut via endoscopy (5, 10, 11, 24), genital tract through genital secretions(12, 25-30) or CNS through cerebrospinal fluid (31-37). Of note, one study in PWH diagnosed with cancer allowed the analysis of HIV reservoirs in selected anatomic tissues collected during autopsy (38, 39). More recently, De Scheerder *et al.* investigated the origins of HIV rebound, and although this study was limited by the number of anatomical sites, they showed that viral rebound originated from diverse cellular and tissue reservoirs (40). To address tissue availability, a peri-mortem observational research cohort, 'The Last Gift', was developed and enrolled PWH diagnosed with a terminal illness from a non-HIV condition. Participants consented to blood sampling before death and donated their bodies for a rapid autopsy after death(41-43). This allowed us to collect ante-mortem blood and post-mortem tissues across 28 anatomical compartments from six participants. Post-mortem tissues were collected within 6 hours of death to maximally preserve tissue and viral integrity(43). In the Last Gift cohort, some participants decided that they no longer wanted to continue their ART in the days and weeks before death. This provided the opportunity to observe and characterize rebounding viral populations in blood and compare them to viral populations found in tissues.

To characterize the HIV reservoirs in collected specimens, we combined new technologies that could deep sequence near full-length (FL) *env* HIV genomes and sensitively quantify HIV DNA(44, 45) with established molecular epidemiology inference methods to assess the viral diversity, divergence, predicted cellular tropism, replication competence, compartmentalization, and migration across the human body.

RESULTS

Cohort. The design of the proposed study is summarized in **Figure 1**. This study evaluated six male participants (see **Supplementary Table S1** and **Supplementary Figure S1**). The median age at the time of death was 57.5 years (range: 52-72). All initiated ART during chronic infection and were virally suppressed for a mean 7.5 years before enrolling into the Last Gift study (range: 3-14 years). Two participants, LG01 and LG04, voluntarily interrupted ART 53 days and ~70 days prior to death respectively. The last HIV RNA blood

plasma level prior to death for LG01 and LG04 was 280 copies/ml and 48,000 copies/ml, respectively. LG03, LG05, LG06 and LG08 had undetectable HIV RNA in blood plasma (<20 copies/ml) until death and at the time of the autopsy. The median CD4 T cell count prior to death was 338/mm³ (range: 174 to 868/mm³).

HIV DNA levels. Overall, HIV DNA levels varied in sampled blood and tissues from ~0 to 658.8 copies/10⁶ cells (median=55.9, IQR:22.9-126.5) (**Supplementary Table 2, Figure 2F and Figure 3**). As expected, HIV DNA levels was lowest in CNS samples varying from a range of 0-4.6 copies/10⁶ cells in the occipital lobe to 1.8-33.6 copies/10⁶ cells in the frontal lobe. Although the sample size was small, there was no discernible difference in HIV DNA levels in tissue between the participants who did and did not interrupt ART ($p>0.05$).

Defective proviruses (HIV DNA). Across all samples from all participants, we recovered 676 single genome FL *env* sequences (**Supplementary Table 2**). Hypermutated and otherwise defective (i.e. non-intact) proviruses (FL *env*) accounted for 10.5% of all sequences (71/676), varying from 0 to 18.3% of all FL *env* sequences in a given tissue (**Supplementary Table 2 and Figure 2B**). Non-intact FL *env* sequences were excluded from further analyses, resulting in 605 evaluated intact FL *env* sequences (55 to 152 sequences per participant, median of 7 sequences /site, interquartile range: [IQR] 5-9). Overall, there was evidence that the proportion of genetically intact FL *env* variants varied between anatomical compartments ($P = 0.016$); however, the pattern of this variation was not consistent across participants ($P = 0.96$). To further investigate this, the 28 compartments were collapsed to 11 semi-related categories (blood plasma, cardiac serum, gut, CNS, kidney, liver, lymphoid, pancreas, PBMC, pericardial adipose, and genital tract), and still no pattern was detected. There was also weak evidence for a difference in the proportion intact FL *env* variants across participants ($P = 0.04$), but this association was weakened further after adjusting for anatomical compartment ($P = 0.27$) and anatomical compartment category ($P = 0.051$). The prevalence of intact FL *env* variants for each participant and anatomical compartment is presented in **Figure 2 and Supplementary Table 2**.

Viral diversity and clonality. Viral diversity of HIV RNA populations in blood and HIV DNA populations in tissues varied between tissues across participants and across tissues within participants (**Figure 2E and Supplementary Table and Figure S2**). In participants who decided to stop therapy, most of the rebounding intact HIV RNA in blood (65% and 80% of variants) was >99% identical, with the mean pairwise nucleotide Tamura Nei 93 distance (46) of 0.015 for LG01, and 0.001 for LG04. Overall, viral diversity did differ significantly across participants ($p=0.057$) but not across tissues ($p=0.51$); however, our power to detect differences across tissues was poor due to the large number of compartments and small number of participants. Considering this, we also looked at individual tissue comparisons to blood plasma HIV RNA populations and found that viral diversity was overall higher for all HIV DNA in tissues but the higher viral diversity was only significant for prostate ($p=0.046$), rectum ($p=0.045$), right colon ($p=0.013$), spleen ($p=0.013$), and occipital cortex ($p=0.006$).

We also found evidence of identical intact FL *env* proviruses within and across tissues. For example, monoclonal FL *env* populations (100% identical) were found in ileum (3 out of 3 intact FL *env* sequences sampled) in LG03, jejunum (7 out of 7) in LG04 and testes (6 out of 6) in LG06. Identical FL *env* sequences

were also found across compartments, with LG03 having 30 identical intact envelopes recovered from 8 anatomical compartments and LG04 having 14 identical intact envelopes from 5 tissue compartments. We also found a large nearly identical population ($\geq 99\%$ identical) of 25 FL *env* proviruses sampled across 10 compartments in LG08 (**Figure 4** and **Supplementary Figure S4**). The proportion of identical FL *env* sequences varied significantly across participants ($p < 0.01$) with the proportion of identical FL *env* variants throughout the body being: LG01 21/75 28%, LG03 (54/113, 48%), LG04 (28%, 42/152), LG05 (14/55, 25%), LG06 (19/104 18%) and LG08 (25/106 24%). While the proportion of identical FL *env* sequences differed significantly across compartments, effect modification was also detected between participant and compartment ($P < 0.001$). In other words, the differences across compartments varied according to participant with no consistent trends of clonality across tissues observed with or without sustained viral suppression.

Viral tropism and re-population. HIV DNA populations in ante-mortem PBMC and post-mortem tissues were predominantly X4-tropic based on genotropism analysis for four participants (LG01 84%, LG03 75.5%, LG05 78.2% and LG06 54.5%), while X4-tropic proviruses comprised only 14.9% and 6.6% sequences obtained from LG04 and LG08 (**Supplementary Table 2** and **Supplementary Figure S5**). Next, we evaluated tissue re-population during rebounding viremia in the 2 participants who voluntarily decided to interrupt their therapy:

Participant LG04 who stopped his ART more than 2 months prior to his death and his last HIV RNA level prior to death in blood plasma was 48,000 copies/ml. The rebounding HIV RNA was 100% intact in *env* and nearly clonal (10 out of 10 intact nearly identical [$\geq 99\%$ identity] FL *env* proviruses) and 80% of variants were R5 tropic (**Supplementary Figures S4 and S5**). Interestingly, identical R5 tropic FL *env* sequences were also found in the HIV DNA population in ante-mortem PBMC (17/23 sequences), and across almost all other sampled tissues collected post-mortem, including lymph nodes, genital tract and gut tissues (**Supplementary Figure S4**). While we cannot be sure of the origin of the rebounding R5-tropic population as ante-mortem samples were only available for PBMC and blood plasma, this observation strongly suggests that many tissue reservoirs were re-populated by R5-tropic virus during rebound.

Participant LG01 stopped his ART 53 days prior to his death and his last HIV RNA level prior to death in blood plasma was 280 copies/ml. Again, the rebounding HIV RNA population was 100% intact (17/17) and predominantly clonal (13 out of 17 intact nearly identical [$\geq 99\%$ identity] FL *env*). Phylogenetic analyses also revealed a large nearly identical FL *env* population of 20 X4 tropic proviruses in post-mortem gut ($n=5$), PBMC ($n=4$) prostate ($n=3$) and blood plasma ($n=8$) (**Figure 4** and **Supplementary Figure S5**).

Viral dispersal and re-population. As expected, all sequences clustered by participant when combined in a single phylogeny (**Supplementary Figure S3**), and using phylogeny-trait (i.e. anatomical site) association measures, we found evidence of viral compartmentalization by anatomical site for all six participants ($P < 0.01$) (**Supplementary Table S3**). Discrete phylogeographic models showed various patterns of viral dispersal (**Figures 4 and 5**, only transition events between locations for which the adjusted BF ≥ 3 are shown).

Among participants who remained virally suppressed (LG03, LG05, LG06 and LG08), the majority of transition events originated from lymph nodes and gut tissues (**Figures 5-6** and **Supplementary Figure S6**).

There was also strong evidence for viral dispersal from lymph nodes to prostate in LG03 (BF = 568) and from rectum to seminal vesicles in LG08 (adjusted BF = 74). Interestingly, there was also positive support for viral transitions from PBMC to frontal lobes in LG05 (adjusted BF = 11) and LG08 (adjusted BF = 5.5).

Among the participants who stopped ART (LG01 and LG04), we sequenced FL *env* from HIV RNA from blood plasma collected ante-mortem during viral rebound (HIV RNA levels of 13,500 copies/mL and 48,000 copies/mL respectively). For LG01, there was strong evidence of viral dispersal from PBMC of all transition events (BFs \geq 20) toward gut sites, prostate, spleen, and blood plasma (adjusted BFs of 449.7, 62, 27 and 22.7 respectively). There was also good support (BFs \geq 10) for transition events from blood plasma toward lymph nodes (adjusted BF = 16.6) and liver (adjusted BF = 10.3) and from liver toward lymph nodes (adjusted BF = 11.8). We also showed positive support for viral migration from frontal lobe toward PBMC (adjusted BF = 3.9) and from lymph nodes toward frontal lobe (adjusted BF = 10.4) (**Figures 5-6**). For LG04, 9 sources of viral migration were strongly-supported (BFs \geq 20), including migration from blood plasma toward gut tissues and intense viral exchanges between the sampled gut sites (**Figure 5 and Supplementary Figure S6**). Interestingly for LG04, we also found evidence of bidirectional transition events within the CNS, from the occipital lobe toward frontal lobe (adjusted BF = 25.4) and from the frontal lobe toward the occipital lobe (adjusted BF = 10.1). The lack of dominance of particular transition types shows that a diverse set of anatomical compartments can act as the source of reservoir virus.

Predictors of viral dispersal. We next used a GLM extension of the discrete phylogeographic model to investigate if the number of proviruses, viral diversity, proportion of X4-tropic viruses, and viral divergence between compartments were associated with potential dispersal rates. We first assessed collinearity between all the factors evaluated, and all correlation coefficients were <0.8 . Consistently for LG03, LG04 and LG06, the model revealed strong evidence that viral migration was associated with limited divergence between the source and recipient compartment (respective adjusted BFs of 389.7, 54.4 and 1690.4, **Supplementary Figure S7**). For LG03 and LG04, there was also evidence of association between the transition rate and higher viral diversity in the recipient compartment (BFs of 373.9 and $\geq 1.5 \cdot 10^4$ respectively). Finally, for LG03 only, there was a positive association between lower proportion of X4 tropic viruses in the source compartment and viral dispersal (BF=5.8). For LG01, LG05 and LG08, there was no evidence that any of the tested factors associated with viral dispersal (**Supplementary Figure S7**).

DISCUSSION

Understanding how HIV populates tissues throughout the human body is crucial for the development of strategies to clear or otherwise lock-down these reservoirs(47-49). Such investigations have been hampered by limitations in technology and access to fresh sampling of tissues throughout the body. This study used samples and data from six participants in a novel, peri-mortem cohort (The Last Gift) of altruistic PWH who provided blood samples before they died and their whole bodies immediately after death. Given the freshness of samples, viral populations likely remained intact, so virologic data were generated and analyzed using state-of-the-art methods. The main findings of this study were: 1) The emergence of large, clonal, intact, HIV RNA populations in blood plasma after stopping ART, and these populations repopulated tissues throughout

the body. 2) Multiple sites can act as hubs for dissemination of HIV within the host, predominately blood, gut and lymph nodes. 3) Viral exchanges occur within CNS areas and between the CNS and blood. 4) Viral dynamics are associated with low HIV divergence between sites and high HIV diversity at the recipient site. Together, these data provide new insights in how HIV populates reservoirs across the human body.

Previous studies showed that lymphoid tissues (e.g. spleen, thymus, lymph nodes, gut-associated lymphoid tissue) are the sites with highest level of viral replication during active infection, and HIV DNA can be readily cultured from lymphoid tissues after years of treatment (5). HIV-infected cells have also been found in the CNS, the lungs, kidneys, liver, adipose tissue, genitourinary tract, and bone marrow (5, 50-53). In agreement with these observations, we found intact FL HIV-1 *env* throughout the 28 anatomic compartments sampled from 6 Last Gift participants. It was interesting that participant LG03 had very high levels of HIV DNA in his pancreas (391.5 gag copies /106 cells [329.2-453.7]), which is also where he had his terminal cancer. Since T cells infiltrate tumor microenvironments, it is possible that the resident tumor attracted HIV-infected T cells(54). More tumor tissues will need to be evaluated to confirm this observation.

Consistent with other reports that investigated blood from PWH(55-57), our study found various proportion of identical FL *env* sequences in PBMC (varying from 19% [CI95%:4-46%] to 81%[CI95%:54-96%] for LG08 and LG01 respectively). We also found identical FL *env* sequences in various tissues (e.g. ileum and jejunum for LG03 and LG04) and across multiple tissues. This observation could be from either monoclonal HIV RNA infecting susceptible cells at multiple sites, as seen in LG04 during rebound viremia, or infected cells that clonally expanded at one site then migrated into a different site. These scenarios are not mutually exclusive. While identical *env* sequences may reflect cellular proliferation, without integration site analysis, we cannot determine if identical *env* sequences are derived from cellular clonal expansion(58). Future integration site analyses are needed to confirm this possibility, as illustrated by Patro *et al* (59).

Concerning viral migration, this study found that not only blood but also mainly gut and lymphoid tissues can act as sources for the dissemination of HIV, which is consistent with previous studies suggesting that both are major HIV reservoirs during ART(60-65). Interestingly, we also saw that genital tract tissues (i.e. prostate, seminal vesicles, testis) were potential sources of viral dispersal to other tissues. Although limited by the number of intact FL *env* proviruses recovered from the frontal and occipital lobes (5 out of 6 participants), this study shows that viral migration occurs within the CNS (LG03 and LG04), between blood and the CNS in 4 participants (LG01, LG03, LG05, LG08) and from the lymph nodes toward the CNS (LG01).

Assessing the factors that may contribute to the viral dynamics within and across compartments is likely to assist in developing strategies targeting the reservoirs; therefore, we used a GLM model to evaluate the role of viral characteristics (e.g. viral diversity, proportion of X4 tropic viruses and viral divergence between compartments) as potential drivers of viral dispersal. Although these findings were limited to 6 participants and only on viral population characteristics, we showed that limited viral divergence between compartment was strongly associated with viral dispersal in 3 out of 6 participants. We also showed that viral dispersal was associated with higher viral diversity in the recipient (2/6 participants) and lower proportion of X4 tropic viruses

in the source (1/6 participants). While these results remain limited and require validation in a larger cohort, it suggests that HIV dynamics across the body is also influenced by molecular characteristics of the HIV populations. Further analyses evaluating pharmacological and immunological factors in the HIV micro-environment would be of interest.

Some PWH at the end of their lives want to stop their ART before they die. This unique situation allowed us to observe viral rebound and tissue re-population in two participants who voluntarily interrupted their ART 53 and 70 days prior to death with rebound viremia. Consistent with De Scheerder *et al* (40), our phylogenetic and statistical analyses show that rebounding viral populations can originate from various anatomical compartments with no prominent source of rebound. Also consistent with previous studies (66), we showed that rebounding HIV RNA populations were comprised of intact nearly identical FL *env* sequences, and thus we found that HIV RNA populations were more homogenous than HIV DNA populations in tissues in the two participants who stopped therapy ($P = 0.025$). Interestingly, when LG04 stopped his ART, he had an asymptomatic rebound of a large, intact, nearly identical R5-tropic HIV RNA population (FL *env*) in blood. This population was identical to a HIV DNA population seen in PBMC ante-mortem. After death, we found identical FL *env* sequences in tissues throughout the body including lymph nodes, genital tract and gut tissues. Similarly, LG01 had rebound viremia with an intact nearly identical X4-tropic population, and we found identical FL *env* sequences in gut, PBMC and prostate. Although not conclusive because we did not sample tissues before stopping ART, this strongly suggest that the rebounding HIV population in blood plasma is the source of viral replenishment in tissue reservoirs. It also supports the notion that preventing HIV rebound in the blood may allow for ART-free remission, like what was observed in Timothy Brown and the London patient, when their previously HIV-susceptible circulating CD4 T cells were replaced by bone marrow transplant with CD4 T cells that were no longer susceptible to HIV infection because of the CCR5 32 Δ deletion (67, 68).

Our study has a number of limitations. The main limitation is the small number of participants, especially having only two participants who stopped their ART. Nonetheless, it allowed us to unprecedentedly observe the population of tissue reservoirs from the blood during rebound viremia. Another limitation is that the participants were all in the process of dying, which may limit the generalizability to healthy PWH. Also, this study, as others(69, 70), focused on the HIV *env* gene, which has the greatest amount of molecular diversity and evolution of all coding regions(71-76), but we acknowledge that by sequencing only the *env* region, we may have incorrectly inferred that some viruses intact in the *env* coding region were replication competent, when they may have had defects in other genome regions. Further, the study found many identical HIV DNA *env* single genomes, consistent with previous reports,(77-80) but we did not directly assess cellular clonal expansion, which is thought to be an important mechanism for HIV persistence(81-83). However, others have shown that multiple identical *env* sequences in proviruses provide a strong indication for clonal expansion(84, 85). Sequencing near FL proviral genome (86-88) would increase the sensitivity of the analyses of intactness and clonality but these approaches require a high cellular input which may limit the ability to explore reservoirs such as the central nervous system. Therefore, we are confident that our approach provided a good surrogate for the extensive analysis of HIV reservoirs. Further, our study did not perform phenotypic testing of CCR5

tropism, and every participant in our study had X4-tropic virus by genotypic analysis. This is likely because participants were infected for a long time before cohort enrollment(89), and emphasizes that people with longstanding HIV infection are unlikely to benefit from measures focusing only on CCR5 tropic viruses, as was the case for the two PWH who received CCR5 mutated bone marrow transplants and have been cured of HIV so far (68, 90). Finally, we cannot exclude the possibility of blood T cell contamination in tissues obtained during autopsy. This is likely to be a small impact on our analysis given the small size of capillaries compared to overall tissue mass and settling of blood in tissues, i.e. livor mortis. While we cannot completely rule out such contamination, our sequence analyses showed viral compartmentalization for all participants, which suggested that possible blood contamination did not significantly impact our analyses.

In conclusion, this study leveraged the unique Last Gift cohort of participants who continued and discontinued ART before dying, after which multiple tissues were rapidly collected at autopsy. From these samples, the study reconstructed the spatial dynamics of HIV across the body and found that multiple body compartments, especially the gut and lymphoid tissues, can act as hubs for dissemination of HIV, and that blood is the likely conduit for dispersal. This study also uniquely clarified the HIV dynamics within the CNS and across the blood brain barrier. Importantly for cure efforts, our analyses strongly indicated that rebounding virus was derived from circulating reservoirs, and that rebounding virus quickly populated deep tissue reservoirs throughout the body via the blood. Given the extent of intact FL env HIV populations throughout the body, it is likely that any of the tissue reservoirs could facilitate re-seeding of virus to the blood and thus repopulate tissue reservoirs. These findings might have consequences for treatment interruption studies although clinical significance is unclear(91). Thus, for cure efforts aimed at eradication, then all such reservoirs would need to be cleared. Alternatively, functional cure efforts could focus on making circulating CD4 T cells impervious to infection to prevent rebound viremia and subsequent viral dispersal and reseed, although this approach would not address local viral reactivation and damage.

METHODS

Study cohort and sampling. Peripheral mononuclear cells (PBMC) and tissue samples were obtained from the first 6 PWH enrolled in 'The Last Gift cohort.(92, 93) Study participants were followed closely near the end of their lives (approximately every other week with closer intervals as death approached). During these visits, participants provided: (i) Detailed clinical and socio-demographic information before their death (use of ART, chemotherapy and other therapies, surgical procedures, co-infections, etc.), (ii) blood collections while they were alive, and (iii) their entire bodies after they died for a rapid autopsy. Clinical characteristics of study participants are summarized in **Supplementary Table S1** and **Supplementary Figure S2**.

Rapid Autopsy. The Last Gift rapid autopsy protocol was designed to collect tissues within 6 hours of death so that post-mortem tissue degradation is minimized. At the time of demise, the body is rapidly transported to the UCSD morgue where the team performs a complete autopsy to sample all organs for formalin-fixed

paraffin embedded tissue histological analysis and snap-frozen tissue collected in liquid nitrogen. Fluids collected include CSF and blood drawn from cardiac puncture.

HIV DNA Quantification and Sequencing.

DNA extraction, purification and quantification. Genomic DNA was extracted from 5 million PBMCs and snap-frozen tissues using QIAamp DNA Mini Kit (Qiagen cat#51306) per manufacturer's protocol. After extraction, precipitation was performed to concentrate DNA. Concentrations of DNA were determined using NanoDrop One (ThermoScientific). Levels of HIV DNA were quantified by droplet digital PCR (ddPCR) from extracted DNA using the BIO-RAD QX200 Droplet Reader.⁽⁹⁴⁾ Copy numbers were calculated as the mean of 3 replicate PCR measurements and normalized to one million cells, as determined by RPP30 (total cell count).^(94, 95)

RNA Extraction and Quantification. RNA was extracted from blood plasma layering 500-700µl of plasma on top of 200µl of 20% sterile filtered sucrose solution. Sample was spun at 23,500xg for 1 hour at 4°C to pellet the virus. Supernatant was removed and the pellet resuspended in 140µl of PBS. RNA was extracted using Qiagen's QIAamp Viral RNA mini kit (cat# 52904) according to the manufacturer's recommendation. cDNA from HIV RNA was generated using Bio-Rad One-Step RT-ddPCR Advanced Kit for Probes (cat# 186-4021), and levels were quantified by ddPCR (BIO-RAD QX200 Droplet Reader). Copy numbers were calculated as the mean of 3 replicates.

Nested PCR. To amplify single genome FL *env*, DNA extracted from antemortem PBMCs and post-mortem tissues was diluted using ddPCR quantification data. This limited dilution PCR reaction can prevent PCR recombination, ambiguous base calls and allow the amplification of viral single genomes.^(87, 96) For HIV RNA in blood plasma, cDNA was generated from RNA using SuperScript III First Strand Synthesis System (cat# 18080-051). Template cDNA and HIV DNA extracted from tissues were diluted until approximately 30% of the second-round reactions were positive for the correctly-sized amplification product. Primers used for the first round were 5'FENVouter (forward) TTAGGCATCTCCTATGGCAGGAA and 3'RENVouter (reverse) TCTTAAAGGTACCTGAGGTCTGACTGG. First round PCRs were performed using the Advantage 2 PCR Kit from Takara (cat# 639206) following manufacturer's recommendations using the 10X SA Buffer. Cycling conditions were 95C for 1 min, 35 cycles of 95C 15sec, 57C 30sec, 68C 3min with a final extension at 68C for 10min. The second round PCR were done using 5'FENVinner: GAGCAGAAGACAGTGGCAATGA (forward) and 3'RENVinner: CCACTTGCCACCCATBTTATAGCA (reverse). Cycling conditions were 95C for 1 min, 30 cycles of 95C 15sec, 64C 30sec, 68C 3min with a final extension at 68C for 10min. PCR clean ups were done on the second round reaction products using QIAquick PCR Purification Kit (cat# 28106). DNA was quantified using Qubit dsDNA HS Assay Kit (Invitrogen cat#Q32854). Quality and integrity were measured using Genomic DNA Screen Tape in combination with the 2200 TapeStation System (Genomic DNA Reagents cat#5067-5366, Genomic DNA Screen Tape Cat #5067-5365).

Nextera XT Library Preparation. Single Genome amplicons were prepared for deep sequencing using the Nextera XT DNA Library Preparation Kit (Illumina FC-131-1096) with indexing of 96-samples per run (Nextera XT index kit set A FC-131-2001) per manufacturer's protocols.

Assembly of full-length HIV env proviruses. We used a custom designed pipeline to recover FL env HIV sequences from the paired end reads. The pipeline included a preliminary step of quality control, which included trimming reads for PHRED quality above or equal 30 and removal of Illumina adapters. Next, overlapping identical paired forward and reverse reads were merged and pre-mapped to HXB2 reference genome. Cleaned reads were re-mapped to the *de novo* assembled near FL env sequence before generating the final consensus sequence. The minimum acceptable coverage was set to 10,000 reads. To identify mixtures (i.e. suggesting multiple amplified HIV templates), we screened all generated FL env contigs. Mixtures were identified based on read coverage and variant calling. Contigs with evidence of single nucleotide polymorphisms (SNPs) with a frequency >1% were considered mixtures and excluded from further analyses.

Test for cross-contamination A maximum likelihood (ML) phylogeny including all sequences from the four participants was estimated using IQtree with the general time reversible (GTR) substitution model (97) to test for contamination, which would show as intermixed clustering of taxa between participants.

Identification of defective or hypermutant sequences. FL envelope containing large deletions (>100bp) were considered defective(86, 98). Deleterious stop codons were identified using the Gene Cutter tool (Los Alamos HIV Database). Any contigs containing a stop codon were considered defective. APOBEC-induced G-A hypermutations were identified using the Los Alamos HIV Database Hypermut 2 program and the participant's consensus sequence(99, 100). Proviruses with a number of mutations significantly higher than the participant's consensus (Fisher exact test $p < 0.05$) were considered hyper-mutant and were not included in the downstream analyses described below.

Sequence analyses.

Identification of Identical full length env sequences. For determining the sequences which were $\geq 99\%$ or 100% genetically identical, we used the ElimDupes tool from Los Alamos HIV database using $\geq 99\%$ and 100% genetic identity threshold as the analysis parameter(101). A sequence was classified as identical if it was a 100% match against another sequences sampled from the same participant. Once identified, the proportion of identical and nearly identical sequences were calculated by dividing the total number of sequences classified as identical/nearly identical for each participant or compartment by the total number of sequences for that group.

Diversity and Divergence. Viral diversity was defined as the average pairwise genetic distance between sequences from a compartment using the Tamura Nei 93 (TN93) correction for multiple hits(46). Viral

divergence was assessed by computing the mean pairwise distance (TN93) between viral populations sampled across anatomical sites.

Co-receptor tropism. Viral tropism of each variant was inferred from the V3 amino acid sequence using geno2pheno(102). We applied a conservative 10% false positive rate threshold for co-receptor CXCR4 usage based on recommendation from the European Consensus Group on clinical management of HIV-1 tropism testing.

Viral migration. Prior to evaluating the within-host migration processes, the level of spatial structure was quantified using the Simmonds association index (AI) implemented in BaTS v1.0(103). We used the BEAST software package v1.10(104) for all evolutionary analysis. All sequences were considered isochronous, i.e. branch lengths were estimated in units of substitutions per site. For this, a strict molecular clock was specified and the clock rate was fixed to 1. The substitution process was described with a HKY+ Γ model(105, 106) and a constant population size was assumed. Discrete trait analyses were performed using the asymmetric diffusion model(107, 108). To identify the subset of migration rates that were most informative to reconstruct the dispersal history we used a model averaging procedure (Bayesian stochastic search variable selection, BSSVS)(107). Bayes factor (BF) support for all possible types of location exchanges were calculated with SpredD3(109). BF between 3-20, 20-150, and above 150 were considered positive, strong, and decisive support respectively(110). Estimates of the posterior expected number of migration events between all pairs of locations (Markov jumps) were computed through stochastic mapping techniques(105, 111). To identify individual-level differences in within-host viral diffusion patterns, the trait analyses were performed on a patient-specific basis.

To investigate whether support for migration links followed from the relative abundance of the involved trait states, analyses were repeated whilst randomly permuting the compartment states between tips during the MCMC sampling(112), a technique analogous to the tip-date randomization procedure for testing the significance of the temporal signal(113, 114). Indeed, if the support for a particular migration rate persisted after randomizing the tip-to-location assignments, one cannot rule out that the support is due to sampling intensity differences. Furthermore, of the migration links that pass the above filter (good BF support in the 'as is' analysis and poor BF support in the 'tip-state-swap' analysis), only those that remain significant after accounting for the sampling heterogeneity were taken into account to further reduce the false positive rate. To this end, the inclusion frequencies from the tip-state-swap analysis instead of those of the 'as is' analysis were treated as the prior inclusion probabilities when recomputing the Bayes factor support, thereafter referred to as "adjusted BF". Only the results with adjusted BF with positive support (BFs \geq 3) were further discussed in this study. This approach was also used to investigate the sensitivity of the Simmonds AI to the sampling heterogeneity.

Multiple MCMC chains were run sufficiently long to ensure convergence and adequate mixing (Effective Sample Size [ESS]>200), which was inspected using Tracer 1.7(115). The chains were sampled every 500,000 generations and combined after removal of the burn-in. Maximum clade credibility (MCC) trees were obtained with TreeAnnotator 1.10(104).

Identifying predictors of within-host spread. The generalized linear model (GLM) extension of the discrete trait model implemented in BEAST(116) was used to investigate the relevance of potential explanatory variables (predictors) to explaining the dispersal across body compartments. The following variables were included: (1) the number of FL intact *env* provirus recovered in each compartment to control for sampling bias effects(116), (2) the level of HIV DNA (ddPCR gag copy/10⁶ cells for all antemortem PBMC and post-mortem tissue samples) or level of HIV RNA for antemortem blood plasma samples (the number of copy/μL), (3) viral diversity estimated using average pairwise genetic distance between each sequences, and (4) the proportion of X4 tropic provirus in each compartment. For each of these variables, the values in the compartment source and recipient were considered. We also included a matrix of pairwise measures of TN93 distance between compartments (i.e. viral divergence). Colinearity between the variables included in the GLM model was evaluated. When 2 variables showed a collinearity coefficient ≥ 0.8 , the analysis was repeated with only one variable in the model.

Statistical analyses. Multivariable logistic regression carried out in R version 3.6.1 using function *glm* and binomial link function, was used to compare the proportion of sequences that were intact and identical (clonal). The independent variables in these analyses were participant and anatomical location. For the continuous average pairwise distance (diversity) outcome, a multiple linear regression was used with assumptions of constant variance and normality of residuals checked and met. Mixed models for both binary and continuous outcomes were analyzed with *glmer* and *lmer* from the *lme4* R library. These mixed models used participant as the grouping factor and included a random term for either intercept alone, or intercept and compartment. Due to the sparse nature of the data, all mixed models had difficulty converging with poor model fits. As such the results and P-values of these models are not reliable and so the results were not presented. Given the large number of tissue compartments, the sparse sampling across participants, and the expected effect modification by participant and compartment, comparisons across specific compartments is not informative.

Study approval. All study participants were at least 18 years of age and provided written, informed consent. Our study (protocol number 160563) was approved by The University of California San Diego (UCSD) Office of Human Research Protections Program. One of the participants exercised his legal right to die option in the state of California(117). No government funds were used as part of the option.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: AC, SG, CI, MP, BV, SD, SR, DS

Performed the experiments: AC, CI, MP

Analyzed the data: AC, SG, DS, BV, SD, TS

Enrolled participants. SG, DS, SR

Wrote and revised the manuscript: AC, SG, DS, MFO, BV, SD, SR, TS

ACKNOWLEDGEMENTS

We are very grateful to the Last Gift participants and their families. We would like to acknowledge the investigators and team members in the CFAR Genomic and Sequencing and Translational Virology Cores. We would like to acknowledge Gemma Caballero, Laura Layman and Brianna Scott for their technical assistance. We would like to acknowledge Steven Hendrix and Susanna Concha Garcia for their assistance in participant enrollment, and Susan Little and Matt Strain for the very helpful discussions.

FUNDING

AC acknowledges funding from the National Institute of Health (AI131971-01, AI036214) and from the University of San Diego Center for AIDS Research (CFAR), a NIH funded program and San Diego CFAR TV Core. SG acknowledges funding from the National Institute of Health (HD094646, AI027763, AI134295, AI68636). DS acknowledges funding from the National Institute of Health (P01 AI31385, P30 AI036214). MP was supported by the San Diego The Veterans Medical Research Foundation (VMRF). S.D. was and B.V. is supported by a postdoctoral fellowship from the Fonds voor Wetenschappelijk Onderzoek (FWO, Belgium). S.D. is currently supported by the Fonds National de la Recherche Scientifique (FNRS, Belgium).

COMPETING FINANCIAL INTERESTS:

AC, SG, SD, BV, MFO, SR, CI, MP, SJL and DMS do not have any commercial or other associations that might pose a conflict of interest.

REFERENCES

1. Whitney JB, Hill AL, Sanisetty S, Penaloza-MacMaster P, Liu J, Shetty M, et al. Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature*. 2014;512(7512):74-7.
2. Cohen MS, Shaw GM, McMichael AJ, and Haynes BF. Acute HIV-1 Infection. *N Engl J Med*. 2011;364(20):1943-54.
3. Mzingwane ML, and Tiemessen CT. Mechanisms of HIV persistence in HIV reservoirs. *Rev Med Virol*. 2017;27(2).
4. Kahn JO, and Walker BD. Acute human immunodeficiency virus type 1 infection. *N Engl J Med*. 1998;339(1):33-9.
5. Wong JK, and Yukl SA. Tissue reservoirs of HIV. *Curr Opin HIV AIDS*. 2016;11(4):362-70.
6. Yukl SA, and J.K. W. In: Hope T, Stevenson M, and Richman D eds. *Encyclopedia of AIDS*. New York: Springer; 2015:1–29.
7. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, 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(6):727-8.
8. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, et al. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*. 1997;278(5341):1295-300.
9. Finzi D, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, et al. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nature medicine*. 1999;5(5):512-7.
10. Yukl SA, Shergill AK, Ho T, Killian M, Girling V, Epling L, et al. The distribution of HIV DNA and RNA in cell subsets differs in gut and blood of HIV-positive patients on ART: implications for viral persistence. *The Journal of infectious diseases*. 2013;208(8):1212-20.
11. van Marle G, Gill MJ, Kolodka D, McManus L, Grant T, and Church DL. Compartmentalization of the gut viral reservoir in HIV-1 infected patients. *Retrovirology*. 2007;4:87.
12. Smith DM, Kingery JD, Wong JK, Ignacio CC, Richman DD, and Little SJ. The prostate as a reservoir for HIV-1. *AIDS (London, England)*. 2004;18(11):1600-2.
13. Shen L, and Siliciano RF. Viral reservoirs, residual viremia, and the potential of highly active antiretroviral therapy to eradicate HIV infection. *J Allergy Clin Immunol*. 2008;122(1):22-8.
14. Saksena NK, and Haddad DeN. Viral reservoirs an impediment to HAART: new strategies to eliminate HIV-1. *Curr Drug Targets Infect Disord*. 2003;3(2):179-206.
15. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim EY, Archer J, Kosakovsky Pond SL, et al. Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature*. 2016.
16. Katlama C, Deeks SG, Autran B, Martinez-Picado J, van Lunzen J, Rouzioux C, et al. Barriers to a cure for HIV: new ways to target and eradicate HIV-1 reservoirs. *Lancet*. 2013;381.
17. Hong FF, and Mellors JW. Changes in HIV reservoirs during long-term antiretroviral therapy. *Curr Opin HIV AIDS*. 2015;10(1):43-8.
18. Hodel F, Patxot M, Snäkä T, and Ciuffi A. HIV-1 latent reservoir: size matters. *Future Virology*. 2016;11(12):785-94.
19. Sengupta S, and Siliciano RF. Targeting the Latent Reservoir for HIV-1. *Immunity*. 2018;48(5):872-95.
20. Archin NM, Sung JM, Garrido C, Soriano-Sarabia N, and Margolis DM. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nat Rev Microbiol*. 2014;12(11):750-64.
21. Archin NM, and Margolis DM. Emerging strategies to deplete the HIV reservoir. *Current opinion in infectious diseases*. 2014;27(1):29-35.
22. Schwartz C, Bouchat S, Marban C, Gautier V, Van Lint C, Rohr O, et al. On the way to find a cure: Purging latent HIV-1 reservoirs. *Biochemical pharmacology*. 2017;146:10-22.
23. Kimata JT, Rice AP, and Wang J. Challenges and strategies for the eradication of the HIV reservoir. *Curr Opin Immunol*. 2016;42:65-70.
24. Yukl S, and Wong JK. Blood and Guts and HIV: Preferential HIV Persistence in GI Mucosa. *J Infect Dis*. 2008;197(5):640-2.
25. Zhu T, Wang N, Carr A, Nam DS, Moor-Jankowski R, Cooper DA, et al. Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: evidence for viral compartmentalization and selection during sexual transmission. *Journal of Virology*. 1996;70(5):3098-107.
26. Wang Y, Frey S, Gupta P, Mittler JE, Nickle DC, Jensen MA, et al.; 2002.

- 9 27. Stekler J, Sycks BJ, Holte S, Maenza J, Stevens CE, Dragavon J, et al. HIV Dynamics in Seminal
0 Plasma during Primary HIV Infection. *AIDS Research and Human Retroviruses*. 2008;24(10):1269-
1 74.
- 2 28. Le Tortorec A, and Dejucq-Rainsford N. [The male genital tract: A host for HIV]. *Gynecologie,*
3 *obstetrique & fertilité*. 2007;35(12):1245-50.
- 4 29. Chaillon A, Smith DM, Vanpouille C, Lisco A, Jordan P, Caballero G, et al. HIV Trafficking Between
5 Blood and Semen During Early Untreated HIV Infection. *J Acquir Immune Defic Syndr*.
6 2017;74(1):95-102.
- 7 30. Chaillon A, Gianella S, Wertheim JO, Richman DD, Mehta SR, and Smith DM. HIV migration
8 between blood and cerebrospinal fluid or semen over time. *The Journal of Infectious Diseases*.
9 2014;209(10):1642-52.
- 0 31. Wang TH, Donaldson YK, Brettelle RP, Bell JE, and Simmonds P. Identification of shared populations
1 of human immunodeficiency virus type 1 infecting microglia and tissue macrophages outside the
2 central nervous system. *Journal of virology*. 2001;75(23):11686-99.
- 3 32. Sturdevant CB, Joseph SB, Schnell G, Price RW, Swanstrom R, and Spudich S.
4 Compartmentalized Replication of R5 T Cell-Tropic HIV-1 in the Central Nervous System Early in
5 the Course of Infection. *PLoS pathogens*. 2015;11(3):e1004720.
- 6 33. Strain MC, Letendre S, Pillai SK, Russell T, Ignacio CC, Günthard HF, et al. Genetic composition of
7 human immunodeficiency virus type 1 in cerebrospinal fluid and blood without treatment and during
8 failing antiretroviral therapy. *Journal of virology*. 2005;79(3):1772-88.
- 9 34. Stam AJ, Nijhuis M, van den Bergh WM, and Wensing AMJ. Differential genotypic evolution of HIV-
0 1 quasispecies in cerebrospinal fluid and plasma: a systematic review. *AIDS Rev*. 2013;15(3):152-
1 61.
- 2 35. Schnell G, Price RW, Swanstrom R, and Spudich S. Compartmentalization and clonal amplification
3 of HIV-1 variants in the cerebrospinal fluid during primary infection. *Journal of virology*.
4 2010;84(5):2395-407.
- 5 36. Oliveira MF, Chaillon A, Nakazawa M, Vargas M, Letendre SL, Strain MC, et al. Early Antiretroviral
6 Therapy Is Associated with Lower HIV DNA Molecular Diversity and Lower Inflammation in
7 Cerebrospinal Fluid but Does Not Prevent the Establishment of Compartmentalized HIV DNA
8 Populations. *PLoS Pathog*. 2017;13(1):e1006112.
- 9 37. Bednar MM, Sturdevant CB, Tompkins LA, Arrildt KT, Dukhovlinova E, Kincer LP, et al.
0 Compartmentalization, Viral Evolution, and Viral Latency of HIV in the CNS. *Curr HIV/AIDS Rep*.
1 2015;12(2):262-71.
- 2 38. Lamers SL, Rose R, Maidji E, Agsalda-Garcia M, Nolan DJ, Fogel GB, et al. HIV DNA Is Frequently
3 Present within Pathologic Tissues Evaluated at Autopsy from Combined Antiretroviral Therapy-
4 Treated Patients with Undetectable Viral Loads. *Journal of Virology*. 2016;90(20):8968.
- 5 39. Rose R, Lamers SL, Nolan DJ, Maidji E, Faria NR, Pybus OG, et al. HIV Maintains an Evolving and
6 Dispersed Population in Multiple Tissues during Suppressive Combined Antiretroviral Therapy in
7 Individuals with Cancer. *J Virol*. 2016;90(20):8984-93.
- 8 40. De Scheerder MA, Vrancken B, Dellicour S, Schlub T, Lee E, Shao W, et al. HIV Rebound Is
9 Predominantly Fueled by Genetically Identical Viral Expansions from Diverse Reservoirs. *Cell Host*
0 *Microbe*. 2019;26(3):347-58.e7.
- 1 41. Vásquez JJ, and Hunt PW. Participating in HIV Cure Research at the End of Life. *Clinical Infectious*
2 *Diseases*. 2019.
- 3 42. Dubé K, Gianella S, Concha-Garcia S, Little SJ, Kaytes A, Taylor J, et al. Ethical considerations for
4 HIV cure-related research at the end of life. *BMC medical ethics*. 2018;19(1):83.
- 5 43. Gianella S, Taylor J, Brown TR, Kaytes A, Achim CL, Moore DJ, et al. Can research at the end of
6 life be a useful tool to advance HIV cure? *Aids*. 2017;31(1):1-4.
- 7 44. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, et al. Defective
8 proviruses rapidly accumulate during acute HIV-1 infection. *Nat Med*. 2016;22(9):1043-9.
- 9 45. Pardons M, Baxter AE, Massanella M, Pagliuzza A, Fromentin R, Dufour C, et al. Single-cell
0 characterization and quantification of translation-competent viral reservoirs in treated and untreated
1 HIV infection. *PLoS Pathog*. 2019;15(2):e1007619.
- 2 46. Tamura K, and Nei M. Estimation of the number of nucleotide substitutions in the control region of
3 mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*. 1993;10(3):512-
4 26.
- 5 47. Chun TW, Moir S, and Fauci AS. HIV reservoirs as obstacles and opportunities for an HIV cure. *Nat*
6 *Immunol*. 2015;16(6):584-9.
- 7 48. van Marle G, Church DL, van der Meer F, and Gill MJ. Combating the HIV reservoirs. *Biotechnology*
8 *& genetic engineering reviews*. 2018;34(1):76-89.

- 9 49. Smith MZ, Wightman F, and Lewin SR. HIV reservoirs and strategies for eradication. *Curr HIV/AIDS Rep.* 2012;9(1):5-15.
- 0 50. Yukl SA, Shergill AK, Ho T, Killian M, Girling V, Epling L, et al. The distribution of HIV DNA and
1 RNA in cell subsets differs in gut and blood of HIV-positive patients on ART: implications for viral
2 persistence. *J Infect Dis.* 2013;208(8):1212-20.
- 3 51. Nolan DJ, Rose R, Rodriguez PH, Salemi M, Singer EJ, Lamers SL, et al. The Spleen Is an HIV-1
4 Sanctuary During Combined Antiretroviral Therapy. *AIDS Res Hum Retroviruses.* 2018;34(1):123-5.
- 5 52. Couturier J, and Lewis DE. HIV Persistence in Adipose Tissue Reservoirs. *Curr HIV/AIDS Rep.*
6 2018;15(1):60-71.
- 7 53. Smith DM, Kingery JD, Wong JK, Ignacio CC, Richman DD, and Little SJ. The prostate as a
8 reservoir for HIV-1. *Aids.* 2004;18(11):1600-2.
- 9 54. Anari F, Ramamurthy C, and Zibelman M. Impact of tumor microenvironment composition on
0 therapeutic responses and clinical outcomes in cancer. *Future Oncol.* 2018;14(14):1409-21.
- 1 55. Cohn LB, Silva IT, Oliveira TY, Rosales RA, Parrish EH, Learn GH, et al. HIV-1 integration
2 landscape during latent and active infection. *Cell.* 2015;160(3):420-32.
- 3 56. Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, et al. HIV latency. Specific HIV integration
4 sites are linked to clonal expansion and persistence of infected cells. *Science.* 2014;345(6193):179-
5 83.
- 6 57. Wagner TA, McLaughlin S, Garg K, Cheung CYK, Larsen BB, Styrchak S, et al. HIV latency.
7 Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection.
8 *Science.* 2014;345(6196):570-3.
- 9 58. Cohn Lillian B, Silva Israel T, Oliveira Thiago Y, Rosales Rafael A, Parrish Erica H, Learn Gerald H,
0 et al. HIV-1 Integration Landscape during Latent and Active Infection. *Cell.* 2015;160(3):420-32.
- 1 59. Patro SC, Brandt LD, Bale MJ, Halvas EK, Joseph KW, Shao W, et al. Combined HIV-1 sequence
2 and integration site analysis informs viral dynamics and allows reconstruction of replicating viral
3 ancestors. *Proceedings of the National Academy of Sciences.* 2019:201910334.
- 4 60. Telwatte S, Lee S, Somsouk M, Hatano H, Baker C, Kaiser P, et al. Gut and blood differ in
5 constitutive blocks to HIV transcription, suggesting tissue-specific differences in the mechanisms
6 that govern HIV latency. *PLOS Pathogens.* 2018;14(11):e1007357.
- 7 61. Yukl S, and Wong JK. Blood and guts and HIV: preferential HIV persistence in GI mucosa. *J Infect*
8 *Dis.* 2008;197(5):640-2.
- 9 62. Embretson J, Zupancic M, Ribas JL, Burke A, Racz P, Tenner-Racz K, et al. Massive covert
0 infection of helper T lymphocytes and macrophages by HIV during the incubation period of AIDS.
1 *Nature.* 1993;362(6418):359-62.
- 2 63. Pantaleo G, Graziosi C, Demarest JF, Butini L, Montroni M, Fox CH, et al. HIV infection is active
3 and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature.*
4 1993;362(6418):355-8.
- 5 64. Wong JK, Gunthard HF, Havlir DV, Zhang ZQ, Haase AT, Ignacio CC, et al. Reduction of HIV-1 in
6 blood and lymph nodes following potent antiretroviral therapy and the virologic correlates of
7 treatment failure. *Proc Natl Acad Sci U S A.* 1997;94(23):12574-9.
- 8 65. Gunthard HF, Havlir DV, Fiscus S, Zhang ZQ, Eron J, Mellors J, et al. Residual human
9 immunodeficiency virus (HIV) Type 1 RNA and DNA in lymph nodes and HIV RNA in genital
0 secretions and in cerebrospinal fluid after suppression of viremia for 2 years. *J Infect Dis.*
1 2001;183(9):1318-27.
- 2 66. Joos B, Fischer M, Kuster H, Pillai SK, Wong JK, Böni J, et al. HIV rebounds from latently infected
3 cells, rather than from continuing low-level replication. *Proceedings of the National Academy of*
4 *Sciences.* 2008;105(43):16725.
- 5 67. Gupta RK, Abdul-Jawad S, McCoy LE, Mok HP, Peppas D, Salgado M, et al. HIV-1 remission
6 following CCR5Delta32/Delta32 haematopoietic stem-cell transplantation. *Nature.*
7 2019;568(7751):244-8.
- 8 68. Hutter G, Nowak D, Mossner M, Ganepola S, Mussig A, Allers K, et al. Long-term control of HIV by
9 CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med.* 2009;360(7):692-8.
- 0 69. Barton K, Hiener B, Winckelmann A, Rasmussen TA, Shao W, Byth K, et al. Broad activation of
1 latent HIV-1 in vivo. *Nat Commun.* 2016;7:12731-.
- 2 70. Lu CL, Pai JA, Nogueira L, Mendoza P, Gruell H, Oliveira TY, et al. Relationship between intact
3 HIV-1 proviruses in circulating CD4(+) T cells and rebound viruses emerging during treatment
4 interruption. *Proc Natl Acad Sci U S A.* 2018;115(48):E11341-e8.
- 5 71. Travers SA, O'Connell MJ, McCormack GP, and McInerney JO. Evidence for heterogeneous
6 selective pressures in the evolution of the env gene in different human immunodeficiency virus type
7 1 subtypes. *J Virol.* 2005;79(3):1836-41.
- 8

- 9 72. Choisy M, Woelk CH, Guegan JF, and Robertson DL. Comparative study of adaptive molecular
0 evolution in different human immunodeficiency virus groups and subtypes. *J Virol*. 2004;78(4):1962-
1 70.
- 2 73. Seibert SA, Howell CY, Hughes MK, and Hughes AL. Natural selection on the gag, pol, and env
3 genes of human immunodeficiency virus 1 (HIV-1). *Mol Biol Evol*. 1995;12(5):803-13.
- 4 74. Holmes EC, Zhang LQ, Simmonds P, Ludlam CA, and Brown AJ. Convergent and divergent
5 sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1
6 within a single infected patient. *Proc Natl Acad Sci U S A*. 1992;89(11):4835-9.
- 7 75. Leitner T, Kumar S, and Albert J. Tempo and mode of nucleotide substitutions in gag and env gene
8 fragments in human immunodeficiency virus type 1 populations with a known transmission history. *J*
9 *Virol*. 1997;71(6):4761-70.
- 0 76. Vrancken B, Alavian SM, Aminy A, Amini-Bavil-Olyae S, and Pourkarim MR. Why comprehensive
1 datasets matter when inferring epidemic links or subgenotyping. *Infection, genetics and evolution :
2 journal of molecular epidemiology and evolutionary genetics in infectious diseases*. 2018;65:350-1.
- 3 77. Schnell G, Price RW, Swanstrom R, and Spudich S. Compartmentalization and clonal amplification
4 of HIV-1 variants in the cerebrospinal fluid during primary infection. *J Virol*. 2010;84(5):2395-407.
- 5 78. Bale MJ, and Kearney MF. Review: HIV-1 phylogeny during suppressive antiretroviral therapy. *Curr*
6 *Opin HIV AIDS*. 2019;14(3):188-93.
- 7 79. Reeves DB, Duke ER, Wagner TA, Palmer SE, Spivak AM, and Schiffer JT. A majority of HIV
8 persistence during antiretroviral therapy is due to infected cell proliferation. *Nat Commun*.
9 2018;9(1):4811-.
- 0 80. Pinzone MR, VanBelzen DJ, Weissman S, Bertuccio MP, Cannon L, Venanzi-Rullo E, et al.
1 Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by
2 clonal expansion. *Nat Commun*. 2019;10(1):728.
- 3 81. Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, et al. The HIV-1
4 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic
5 changes over time. *Proceedings of the National Academy of Sciences of the United States of*
6 *America*. 2013;110(51):E4987-E96.
- 7 82. von Stockenstrom S, Odevall L, Lee E, Sinclair E, Bacchetti P, Killian M, et al. Longitudinal Genetic
8 Characterization Reveals That Cell Proliferation Maintains a Persistent HIV Type 1 DNA Pool
9 During Effective HIV Therapy. *J Infect Dis*. 2015;212(4):596-607.
- 0 83. Wagner TA, McKernan JL, Tobin NH, Tapia KA, Mullins JI, and Frenkel LM. An increasing
1 proportion of monotypic HIV-1 DNA sequences during antiretroviral treatment suggests proliferation
2 of HIV-infected cells. *J Virol*. 2013;87(3):1770-8.
- 3 84. Hosmane NN, Kwon KJ, Bruner KM, Capoferri AA, Beg S, Rosenbloom DI, et al. Proliferation of
4 latently infected CD4(+) T cells carrying replication-competent HIV-1: Potential role in latent
5 reservoir dynamics. *J Exp Med*. 2017;214(4):959-72.
- 6 85. Laskey SB, Pohlmeier CW, Bruner KM, and Siliciano RF. Evaluating Clonal Expansion of HIV-
7 Infected Cells: Optimization of PCR Strategies to Predict Clonality. *PLoS Pathog*.
8 2016;12(8):e1005689.
- 9 86. Hiener B, Horsburgh BA, Eden JS, Barton K, Schlub TE, Lee E, et al. Identification of Genetically
0 Intact HIV-1 Proviruses in Specific CD4(+) T Cells from Effectively Treated Participants. *Cell*
1 *reports*. 2017;21(3):813-22.
- 2 87. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, et al. Defective
3 proviruses rapidly accumulate during acute HIV-1 infection. *Nature medicine*. 2016;22(9):1043-9.
- 4 88. Lee GQ, Orlova-Fink N, Einkauf K, Chowdhury FZ, Sun X, Harrington S, et al. Clonal expansion of
5 genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. *J Clin Invest*. 2017;127(7):2689-96.
- 6 89. Weinberger AD, and Perelson AS. Persistence and emergence of X4 virus in HIV infection.
7 *Mathematical biosciences and engineering : MBE*. 2011;8(2):605-26.
- 8 90. Allers K, Hutter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, et al. Evidence for the cure of
9 HIV infection by CCR5Delta32/Delta32 stem cell transplantation. *Blood*. 2011;117(10):2791-9.
- 0 91. Julg B, Dee L, Ananworanich J, Barouch DH, Bar K, Caskey M, et al. Recommendations for
1 analytical antiretroviral treatment interruptions in HIV research trials-report of a consensus meeting.
2 *The lancet HIV*. 2019;6(4):e259-e68.
- 3 92. Quinn GP, Murphy D, Pratt C, Munoz-Antonia T, Guerra L, Schabath MB, et al. Altruism in terminal
4 cancer patients and rapid tissue donation program: does the theory apply? *Med Health Care Philos*.
5 2013;16(4):857-64.
- 6 93. Lintz KC, Penson RT, Chabner BA, Mack S, and Lynch TJ. Schwartz Center Rounds. A Staff
7 Dialogue on Phase I Trials: Psychosocial Issues Faced by Patients, Their Families, and Caregivers.
8 *Oncologist*. 1998;3(5):357-64.

- 9 94. Strain MC, Lada SM, Luong T, Rought SE, Gianella S, Terry VH, et al. Highly precise measurement
0 of HIV DNA by droplet digital PCR. *PLoS ONE*. 2013;8(4).
- 1 95. Pinheiro LB, Coleman VA, Hindson CM, Herrmann J, Hindson BJ, Bhat S, et al. Evaluation of a
2 droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem*.
3 2012;84(2):1003-11.
- 4 96. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, et al. Replication-competent
5 noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell*. 2013;155(3):540-
6 51.
- 7 97. Schmidt HA, Minh BQ, von Haeseler A, and Nguyen L-T. IQ-TREE: A Fast and Effective Stochastic
8 Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*.
9 2014;32(1):268-74.
- 0 98. Pinzone MR, VanBelzen DJ, Weissman S, Bertuccio MP, Cannon L, Venanzi-Rullo E, et al.
1 Longitudinal HIV sequencing reveals reservoir expression leading to decay which is obscured by
2 clonal expansion. *Nat Commun*. 2019;10(1):728.
- 3 99. Rose PP, and Korber BT. Detecting hypermutations in viral sequences with an emphasis on G --> A
4 hypermutation. *Bioinformatics*. 2000;16(4):400-1.
- 5 100. Los Alamos National Laboratory, and National Institutes of Health. HIV databases.
6 <http://www.hiv.lanl.gov/>. 2019.
- 7 101. HIV Los Alamos National Laboratory. ElimDupes. [https://www.hiv.lanl.gov/content/sequence/
8 ELIMDUPES/ElimDupesExplain.html](https://www.hiv.lanl.gov/content/sequence/ELIMDUPES/ElimDupesExplain.html).
- 9 102. Beerwinkel N, Däumer M, Oette M, Korn K, Hoffmann D, Kaiser R, et al. Geno2pheno: Estimating
0 phenotypic drug resistance from HIV-1 genotypes. *Nucleic Acids Research*. 2003;31(13):3850-5.
- 1 103. Parker J, Rambaut A, and Pybus OG. Correlating viral phenotypes with phylogeny: accounting for
2 phylogenetic uncertainty. *Infection, genetics and evolution : journal of molecular epidemiology and
3 evolutionary genetics in infectious diseases*. 2008;8(3):239-46.
- 4 104. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, and Rambaut A. Bayesian phylogenetic
5 and phylodynamic data integration using BEAST 1.10. *Virus Evol*. 2018;4(1):vey016.
- 6 105. Minin VN, Bloomquist EW, and Suchard MA. Smooth skyride through a rough skyline: Bayesian
7 coalescent-based inference of population dynamics. *Molecular Biology and Evolution*.
8 2008;25(7):1459-71.
- 9 106. Drummond AJ, Ho SYW, Phillips MJ, and Rambaut A. Relaxed phylogenetics and dating with
0 confidence. *PLoS Biology*. 2006;4(5).
- 1 107. Lemey P, Rambaut A, Drummond AJ, and Suchard MA. Bayesian phylogeography finds its roots.
2 *PLoS Computational Biology*. 2009;5(9).
- 3 108. Edwards CJ, Suchard MA, Lemey P, Welch JJ, Barnes I, Fulton TL, et al. Ancient hybridization and
4 an Irish origin for the modern polar bear matriline. *Current biology : CB*. 2011;21(15):1251-8.
- 5 109. Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, and Lemey P. Spread3: Interactive
6 Visualization of Spatiotemporal History and Trait Evolutionary Processes. *Mol Biol Evol*.
7 2016;33(8):2167-9.
- 8 110. Kass RE, and Raftery AE. Bayes Factors. *Journal of the American Statistical Association*.
9 1995;90(430):773-95.
- 0 111. Minin VN, and Suchard MA. Counting labeled transitions in continuous-time Markov models of
1 evolution. *J Math Biol*. 2008;56(3):391-412.
- 2 112. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. HIV epidemiology. The
3 early spread and epidemic ignition of HIV-1 in human populations. *Science*. 2014;346(6205):56-61.
- 4 113. Firth C, Kitchen A, Shapiro B, Suchard MA, Holmes EC, and Rambaut A. Using time-structured data
5 to estimate evolutionary rates of double-stranded DNA viruses. *Molecular Biology and Evolution*.
6 2010;27(9):2038-51.
- 7 114. Trovão NS, Baele G, Vrancken B, Bielejec F, Suchard MA, Fargette D, et al. Host ecology
8 determines the dispersal patterns of a plant virus. *Virus evolution*. 2015;1(1):vev016-vev.
- 9 115. Rambaut A, Drummond AJ, Xie D, Baele G, and Suchard MA. Posterior Summarization in Bayesian
0 Phylogenetics Using Tracer 1.7. *Syst Biol*. 2018;67(5):901-4.
- 1 116. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying viral genetics and
2 human transportation data to predict the global transmission dynamics of human influenza H3N2.
3 *PLoS Pathog*. 2014;10(2):e1003932.
- 4 117. Danielson MM, and Dubé K. Michael's Testimonial. *Annals of Internal Medicine*. 2018;169(5):349-
5

FIGURE 1. Flow Diagram. Six participants enrolled in the Last Gift cohort (n=2 who stopped ART and n=4 who remained virally suppressed on ART until death) were included in this study. (1) Pre-mortem blood plasma (n=2 participants) and peripheral blood mononuclear cell (PBMC) samples (n=6 participants) were collected before death and tissues were collected during the rapid autopsy procedure. (2) HIV RNA and DNA were extracted from blood plasma and PBMC/tissues for quantification of HIV DNA/RNA (digital droplet PCR); (3) and (4) HIV full length envelopes were sequenced via single genome amplification and sequencing. Intact full length *env* sequences from all samples were used to (5) characterize the HIV populations within each compartments and in blood, (6) assess viral dispersal across tissues using Bayesian phylodynamic models and (7) evaluate factors associated with viral dispersal.

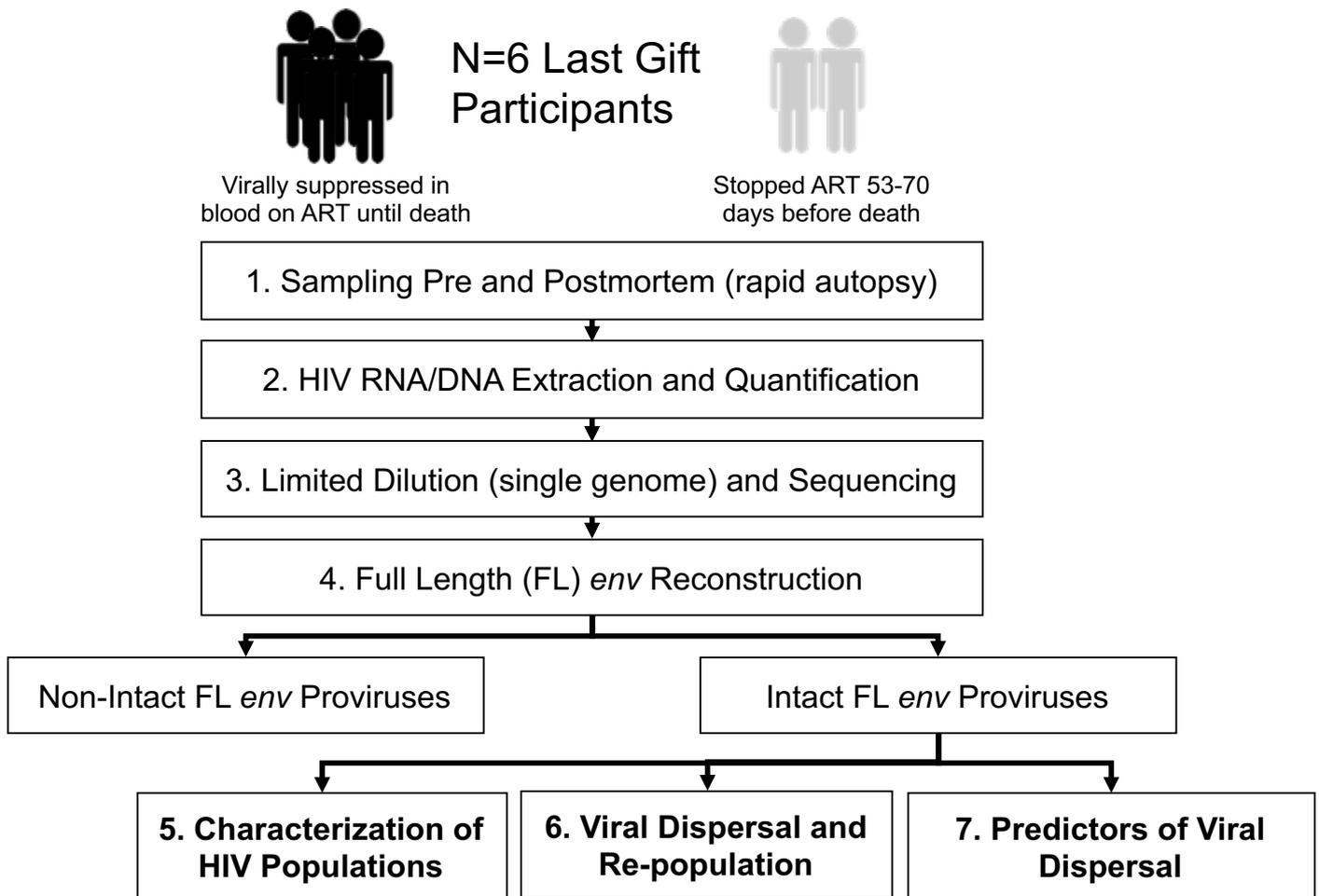


FIGURE 2. Characteristics of the HIV Populations Within Each Compartments. Scatter dot plots sorted by sampled compartment and colored by participants with the y-axis showing respectively. **1A.** The total N of sequences; **2B.** The proportion of intact; **2C.** The proportion of nearly identical (any sequences with $\geq 99\%$ nucleotide identity with at least one other sequence from the same participant); **2D.** Proportion of identical (100%); **2E.** Pairwise genetic distance between sequences and **2F.** levels of HIV DNA. The pairwise genetic distance between sequences from a compartment was measured using the Tamura Nei 93 (TN93) algorithm (46). Level of HIV DNA was quantified by droplet digital PCR (ddPCR). Copy numbers were calculated as the mean of 3 replicate PCR measurements and normalized to one million cells, as determined by RPP30 (total cell count)(94, 95). See method section for details.

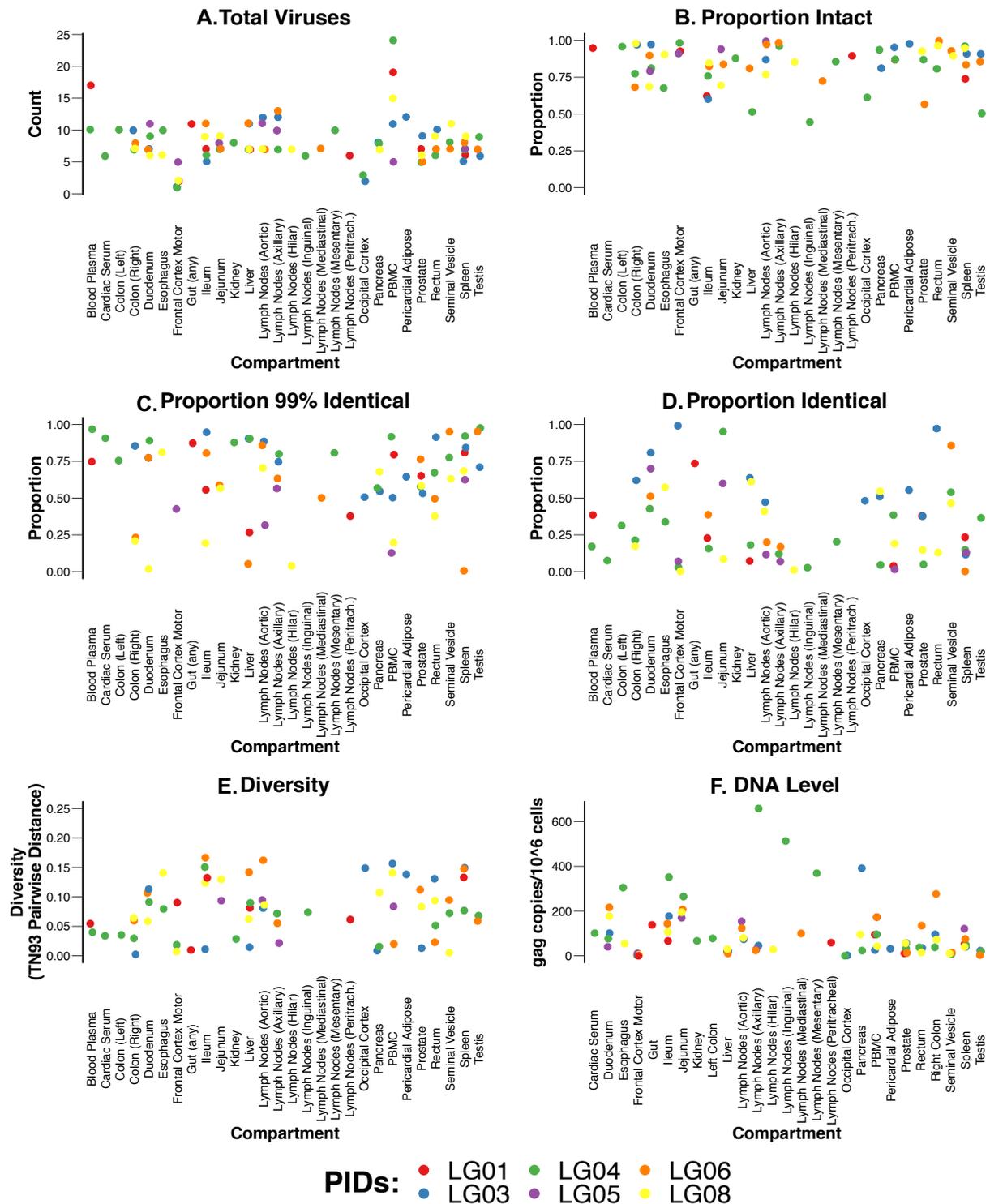


FIGURE 3. HIV DNA Levels in Compartments for each Participant. Level of HIV DNA was quantified by droplet digital PCR (ddPCR). Copy numbers were calculated as the mean of 3 replicate PCR measurements and normalized to one million cells, as determined by RPP30 (total cell count)(94, 95). See method section for details. The dots and horizontal bars represent the mean and range (minimum and maximum) of the HIV DNA levels.

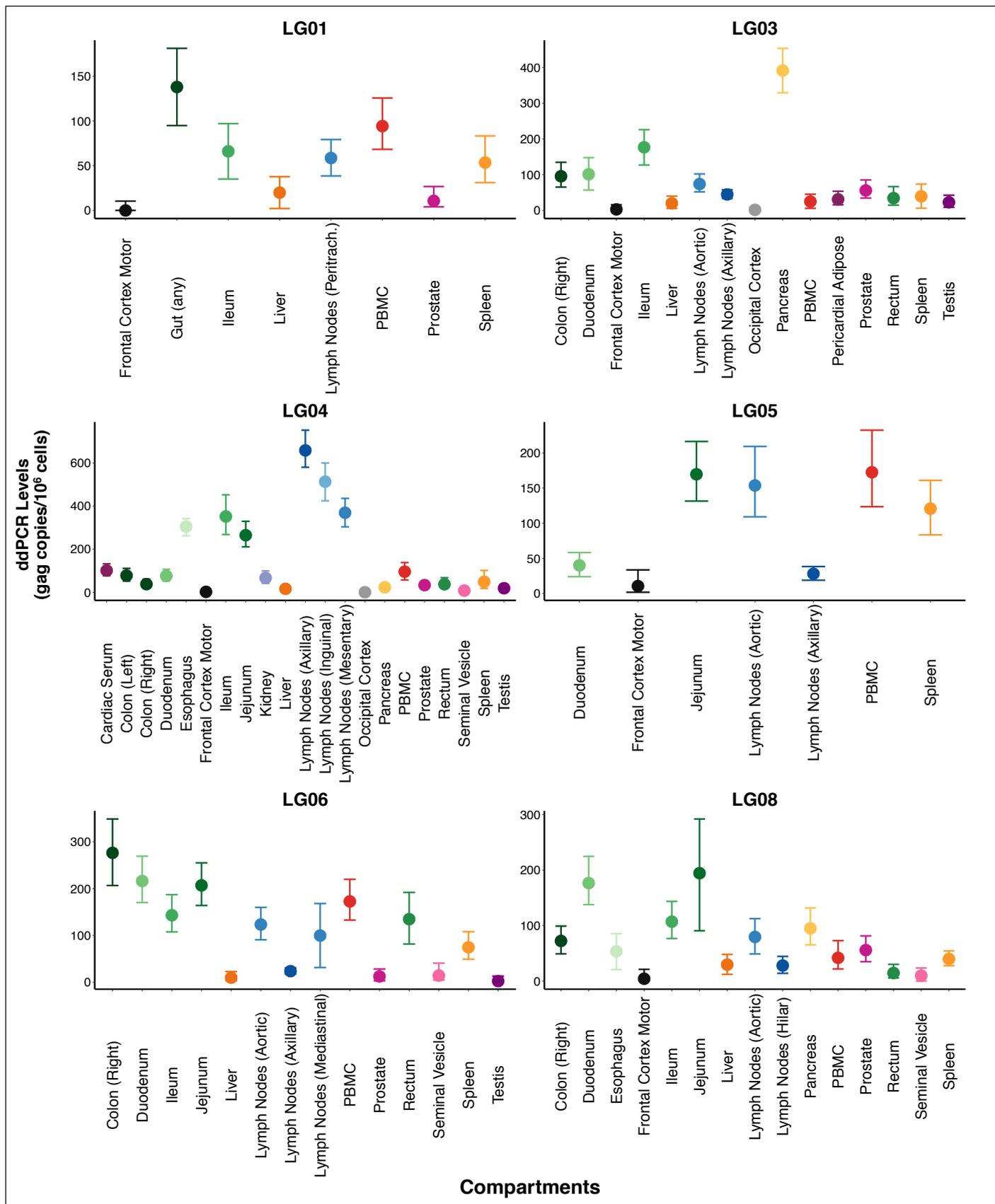
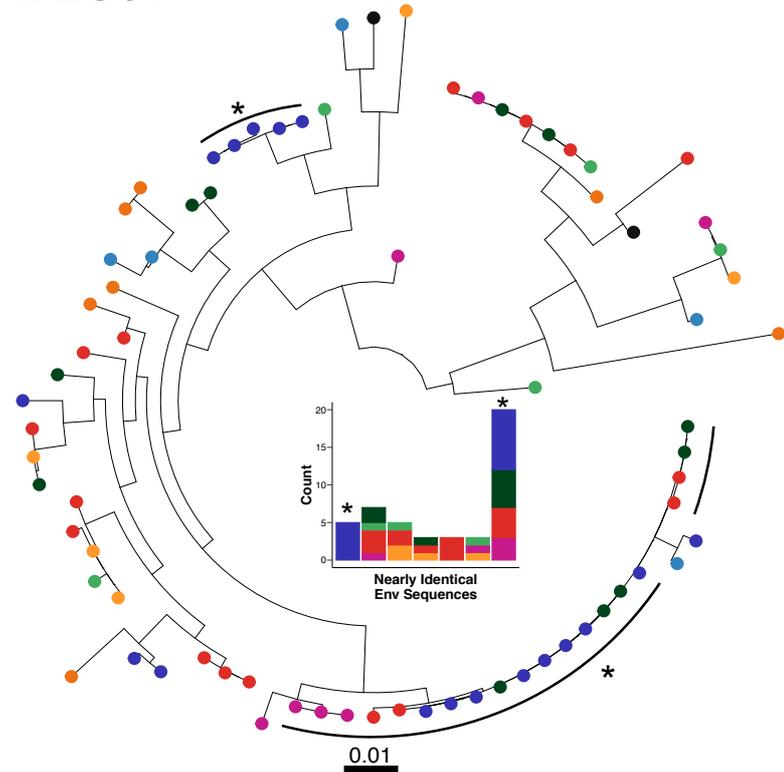


FIGURE 4. Maximum Likelihood (ML) Phylogenies and Clonal Populations (Full Length Envelope) for 2 Participants Who Stopped ART (A) or Remained Virally Suppressed on ART (B) . Phylogenies were estimated using IQtree(97) from the full length HIV *env* sequences obtained from pre-mortem blood plasma and from tissues and PBMC collected during rapid autopsy. **A.** ML phylogeny for Individual LG01 who stopped therapy. **B.** ML phylogeny for Individual LG03 who remained virally suppressed. Tips are colored by compartment as in the legend. Size and distribution of nearly clonal FL *env* populations (99% identical, populations of at least 3 identical proviruses) for each participant are presented in the middle of each tree. Colors represent tissues described in Left Panel legend. For LG01, nearly identical populations including HIV RNA viruses sampled in blood plasma during viral rebound are marked with *. See also **Supplementary Figure S4** for the 4 remaining participants.

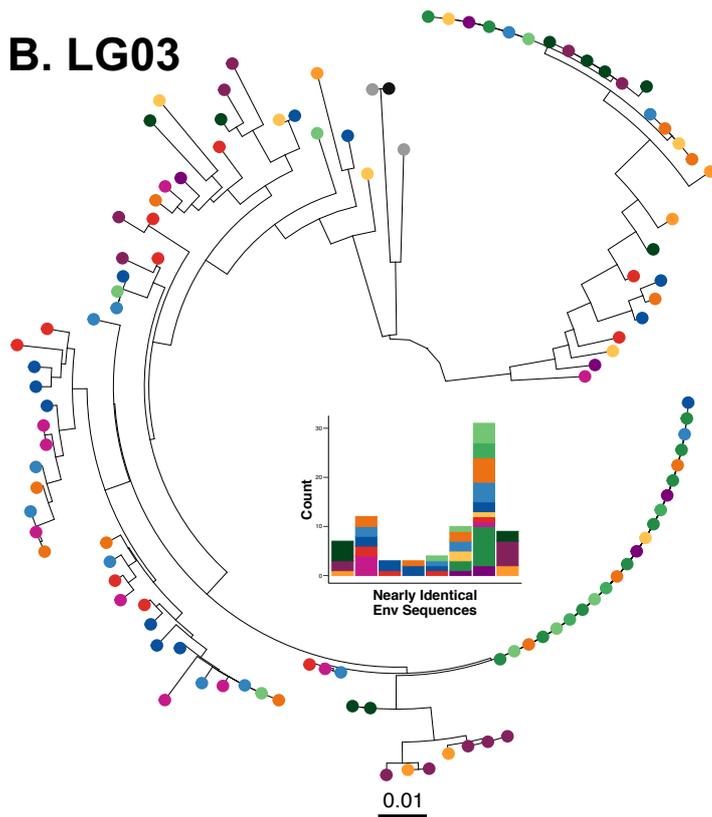
Compartment:

- Blood Plasma
- Cardiac Serum
- Colon (Left)
- Colon (Right)
- Duodenum
- Esophagus
- Frontal Cortex Motor
- Gut (any)
- Ileum
- Jejunum
- Kidney
- Liver
- Lymph Nodes (Aortic)
- Lymph Nodes (Axillary)
- Lymph Nodes (Hilar)
- Lymph Nodes (Inguinal)
- Lymph Nodes (Mediastinal)
- Lymph Nodes (Mesentary)
- Lymph Nodes (Peritrach.)
- Occipital Cortex
- Pancreas
- PBMC
- Pericardial Adipose
- Prostate
- Rectum
- Seminal Vesicle
- Spleen
- Testis

A. LG01

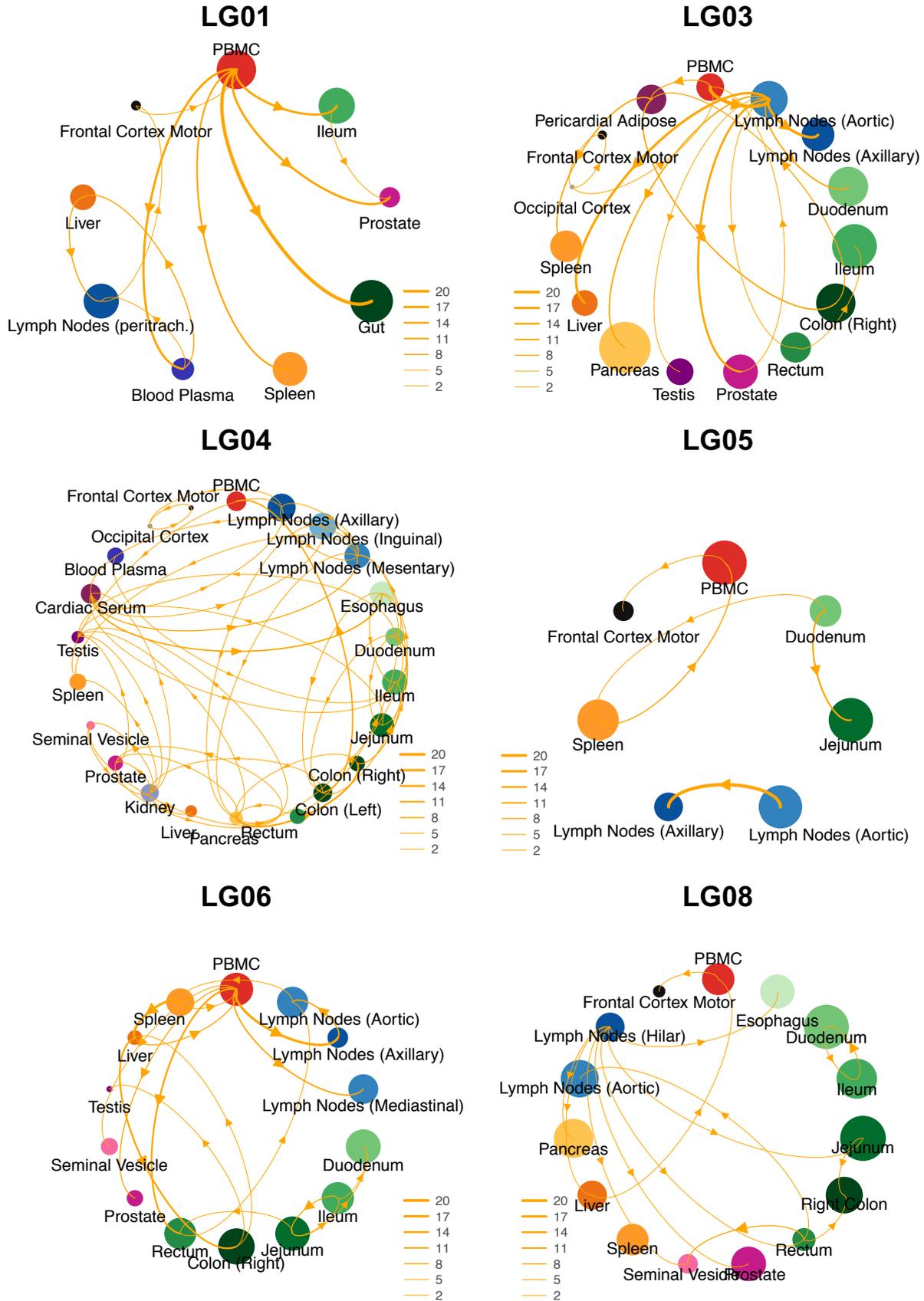


B. LG03



1
2
3
4
5

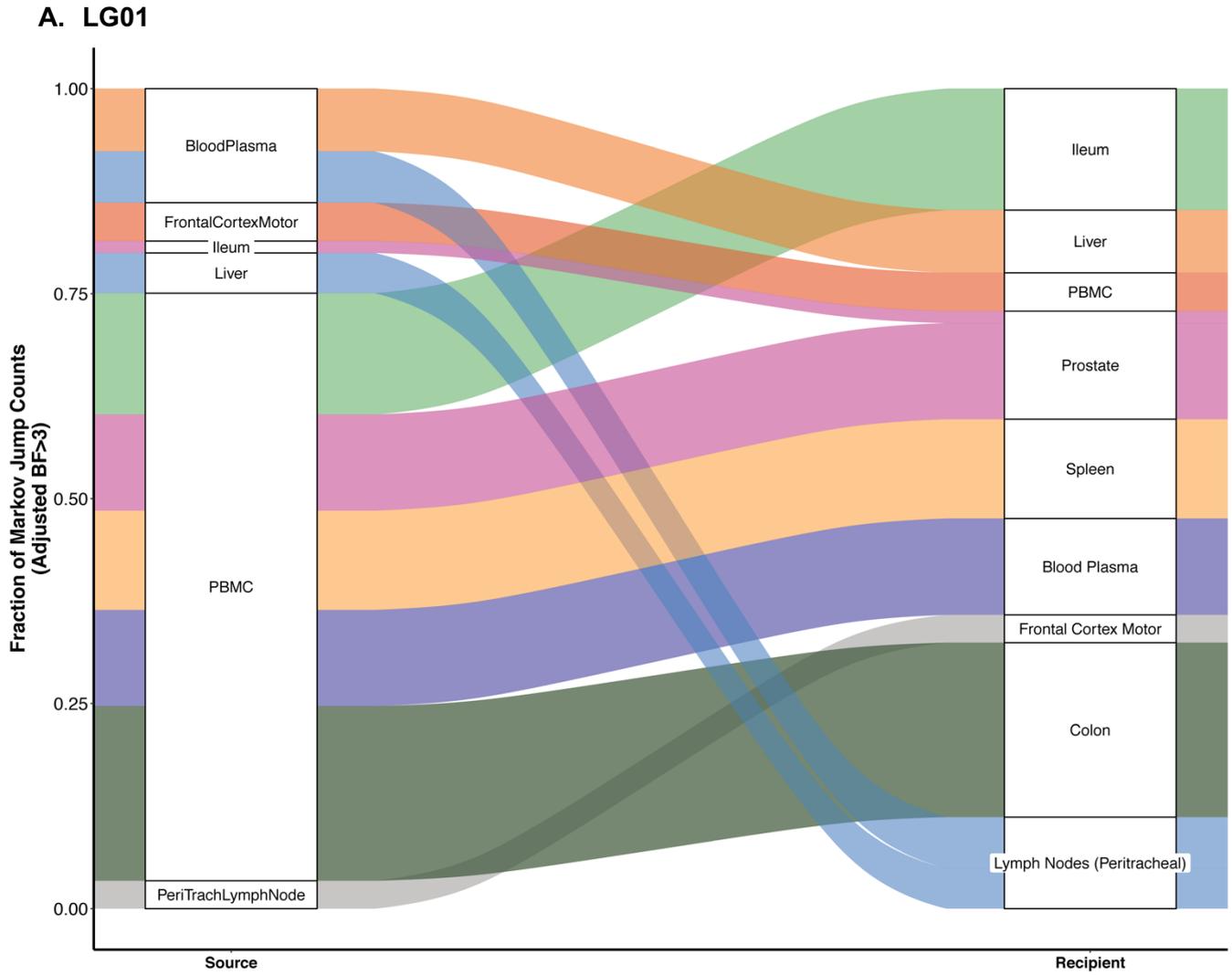
FIGURE 5. Lineage Dispersal Events between Compartments. Circle size is proportional to HIV DNA level (or RNA for plasma) in each compartment. The thickness of the arrows corresponds to the average number of inferred migration events between compartments. Only transition events between locations for which the adjusted BF ≥ 3 (at least positive evidence) are shown.



6

7 **FIGURE 6. Proportion of Supported Transition Events between Compartments for 2 Participants**
 8 **Who Stopped ART (A) or Remained Virally Suppressed on ART (B)** . Sankey plot showing the
 9 proportion of transition events between locations for which the adjusted BF ≥ 3 (at least positive evidence).
 0 The adjusted BF support for each transition type is given next to the corresponding color. The source
 1 locations are depicted on the left side of the plots while the destination locations are given at the right side.
 2 See also **Supplementary Figure S6**.

3
4



5
6

B. LG03

