Impact of Antiretroviral Therapy Duration on HIV-1 Infection of T Cells within Anatomic Sites

Eunok Lee,[a] Susanne von Stockenstrom,[b*] Vincent Morcilla,[a] Lina Odevall,[b*] Bonnie Hiener,[a] Wei Shao,[c] Wendy Hartogensis,[d*] Peter Bacchetti,[e] Jeffrey Milush,[d] Teri Liegler,[d] Elizabeth Sinclair,[d] Hiroyu Hatano,[d*] Rebecca Hoh,[d] Ma Somsouk,[d] Peter Hunt,[d] Eli Boritz,[f] Daniel Douek,[f] Remi Fromentin,[g] Nicolas Chomont,[g] Steven G. Deeks,[d] Frederick M. Hecht,[d] Sarah Palmer[a]

[a]The Westmead Institute for Medical Research, University of Sydney, Westmead, New South Wales, Australia

[b]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden

[c]Advanced Biomedical Computing Center, Leidos Biomedical Research Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA

[d]Department of Medicine, University of California San Francisco, San Francisco, California, USA

[e]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

[f]Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

[g]Centre de Recherche du CHUM and Department of Microbiology, Infectiology and Immunology, Université de Montréal, Montreal, Canada

ABSTRACT Understanding the impact of antiretroviral therapy (ART) duration on HIV-infected cells is critical for developing successful curative strategies. To address this issue, we conducted a cross-sectional/inter-participant genetic characterization of HIV-1 RNA from pre- and on-therapy plasmas and HIV-1 DNA from CD4$^+$ T cell subsets derived from peripheral blood (PB), lymph node (LN), and gut tissues of 26 participants after 3 to 17.8 years of ART. Our studies revealed in four acute/early participants who had paired PB and LN samples a substantial reduction in the proportion of HIV-infected cells per year on therapy within the LN. Extrapolation to all 12 acute/early participants estimated a much smaller reduction in the proportion of HIV-1-infected cells within LNs per year on therapy that was similar to that in the participants treated during chronic infection. LN-derived effector memory T (T_{EM}) cells contained HIV-1 DNA that was genetically identical to viral sequences derived from pre- and on-therapy plasma samples. The proportion of identical HIV-1 DNA sequences increased within PB-derived T_{EM} cells. However, the infection frequency of T_{EM} cells in PB was stable, indicating that cellular proliferation that compensates for T cell loss over time contributes to HIV-1 persistence. This study suggests that ART reduces HIV-infected T cells and that clonal expansion of HIV-infected cells maintains viral persistence. Importantly, LN-derived T_{EM} cells are a probable source of HIV-1 genomes capable of producing infectious HIV-1 and should be targeted by future curative strategies.

IMPORTANCE HIV-1 persists as an integrated genome in CD4$^+$ memory T cells during effective therapy, and cessation of current treatments results in resumption of viral replication. To date, the impact of antiretroviral therapy duration on HIV-infected CD4$^+$ T cells and the mechanisms of viral persistence in different anatomic sites is not clearly elucidated. In the current study, we found that treatment duration was associated with a reduction in HIV-infected T cells. Our genetic analyses revealed that CD4$^+$ effector memory T (T_{EM}) cells derived from the lymph node appeared to contain provirus that was genetically identical to plasma-derived virions. Moreover, we found that cellular proliferation counterbalanced the decay of HIV-infected cells throughout therapy. The contribution of cellular proliferation to viral persistence is particularly significant in T_{EM} cells. Our study emphasizes the importance of HIV-1 intervention and provides new in-

sights into the location of memory T cells infected with HIV-1 DNA, which is capable of contributing to viremia.

Antiretroviral therapy (ART) effectively suppresses but does not eradicate HIV-1 infection. HIV-1 persists in a quiescent state within cellular and tissue reservoirs of HIV-1-infected individuals on effective ART (1–7). Although ART can effectively halt HIV-1 replication, the virus persists in plasma and cellular reservoirs, and treatment cessation invariably results in the resumption of viral replication (8–11). Studies have shown that HIV-1 infection of resting CD4$^+$ T cells happens when an activated CD4$^+$ T cell becomes infected by HIV-1 but transitions to a quiescent memory state before the infection eliminates the cell, and it has been proposed that HIV-1 has the potential to directly infect resting CD4$^+$ T cells (4, 7, 12–16).

The number of HIV-1-infected cells is remarkably stable during ART (7, 17). The mechanisms driving the stability of HIV-1 in cellular and tissue reservoirs during suppressive ART are under investigation. The replenishment of these infected cells by low-level ongoing viral replication in anatomic sites containing suboptimal levels of antiviral drugs has been proposed (18–22). However, studies of participants on long-term ART have failed to detect HIV-1 evolution, indicating that viral replication during effective ART is limited (23, 24). The presence of identical HIV-1 DNA sequences among diverse viral genetic populations indicates that these persistent HIV-1-infected cells are partially maintained by cellular proliferation (17, 23, 25–28). Such cellular mechanisms may be major contributors to the persistence of HIV-1-infected cells and the preservation of proviral HIV-1 DNA during suppressive ART.

The majority of published studies have focused on CD4$^+$ T cell subsets within the peripheral blood (PB) when characterizing persistent HIV-1 during ART (29). However, the lymph node (LN) and gut are preferential sites for HIV-1 persistence during therapy (30–33). Moreover, it has been found that HIV-1 can utilize a broad range of coreceptors, including CCR6 and CXCR5, indicating that the virus can establish infection in T cells expressing these markers (34, 35). Therefore, our study aims to define the impact of ART duration on the dynamics and genetic composition of the *gag-pol* region (p6 through nucleotides 1 to 900 of the gene encoding reverse transcriptase [p6-RT]) of HIV-1 within a broad range of T cell subsets derived from different anatomic sites. We performed cross-sectional/interparticipant analysis of HIV-1 DNA sequences in CD4$^+$ T cell subsets derived from the peripheral blood, lymph node, and gut tissues of 26 participants who had received 3 to 17.8 years of suppressive ART. We modeled the impact of therapy duration on the proportion of HIV-1-infected cells and the genetic nature of the virus to understand the cellular mechanisms contributing to viral persistence during therapy. Moreover, we genetically compared HIV-1 RNA sequences derived from pretherapy and early on-therapy plasma and viral DNA sequences derived from CD4$^+$ T cell subsets sorted from the anatomic sites to identify intracellular HIV-1 sources contributing to viremia during ART.

Our study suggested a decline in the proportion of T cells that were HIV-1 infected. We found no substantial accumulation of genetically defective HIV-1 sequences in participants who initiated ART during acute/early and chronic infection, which indicates that the pool of defective viral genomes is established in cells during multiple rounds of HIV-1 replication before viral suppression. Moreover, the genetic comparison of viral populations between plasma and a broad spectrum of CD4$^+$ T cell subsets indicated that lymph node-derived CD4$^+$ effector memory T (T$_{EM}$) cells are a likely source of HIV-1 genomes capable of producing infectious virus. Furthermore, our in-depth genetic analysis revealed that cellular proliferation contributes to HIV-1 persistence by restoring the overall stability of HIV-1-infected cells despite T cell loss during therapy.

RESULTS

HIV-1 infection frequencies of T cells located in different anatomic sites during effective ART. The impact of ART duration on the proportion of HIV-1-infected T cells is not clearly defined. To evaluate the effect of ART duration on the proportion of infected T cells, we performed a cross-sectional/interparticipant analysis of the proportion of HIV-1-infected cells in CD4$^+$ T cell subsets sorted from PB, LN, and gut tissues. We sorted a broad range of CD4$^+$ T cell subsets from the anatomic sites using their specific cellular markers in 26 participants after they had been on effective ART for 3.0 to 17.8 years: 12 who initiated therapy during acute/early HIV-1 infection ($\leq$6 months of infection before initiation of therapy) (AHI group) and 14 who initiated therapy during chronic HIV-1 infection ($\geq$1 year of infection before initiation of therapy) (CHI group) (Tables 1 to 3 and Fig. 1 to 3). The anatomic regions and cellular subsets were collected after the stated duration of ART for each participant (Table 1). These participants were continuously suppressed during the study, except for one participant who had a viral rebound at the time of sampling (Table 1). We used a previously described maximum-likelihood method to estimate the proportion of cells infected within each T cell subset and within each anatomic site (23). The influence of ART duration on the proportions of infected cells are presented as a fold effect per year of ART. We estimated the fold differences in the proportions of cells infected between earlier and later time points. A fold effect per year of ART greater than 1 indicates that a higher proportion of HIV-1-infected T cells is associated with each additional year on therapy, while a value of 1 means a stable proportion of infected T cells during treatment. A fold effect per year of less than 1 indicates that a lower proportion of infected T cells is associated with each additional year of ART. The statistical significance indicates the evidence for increase (a fold effect per year of greater than 1) or decrease (a fold effect per year of less than 1) versus the null hypothesis of no change over the duration of ART (a fold effect per year of 1). Furthermore, we estimated the impact of ART duration on the proportions of infected cells when at least four participants contributed to the fold effect per year on ART within each T cell subset and tissue.

In PB-derived T cells from the AHI group, each additional year of ART was associated with an overall lower proportion of HIV-1-infected T cells that was statistically significant (fold effect = 0.82/year; 95% confidence interval [CI] = 0.70 to 0.97; P = 0.034) (Fig. 4A). This equates to an 18% reduction in the proportion of cells infected during each additional year on therapy. The fold effect on the proportion of cells infected per year in naive (T_N), central memory (T_{CM}), transitional memory (T_{TM}), and effector memory (T_{EM}) T cell subsets sorted from the AHI group ranged from 0.71 to 0.90, all indicating that a lower proportion of infected T cells was associated with each additional year of therapy. This lower proportion of HIV-1-infected T cells was statistically significant in T_N and T_{CM} cells. As we obtained stem cell memory T (T_{SCM}) cells from only two participants from the AHI group, we were not able to examine the association between the ART duration and the proportion of infected T cells within this T cell subset (Table 2). For the CHI group, the fold effect per year of ART within the infected T cells from the PB was similar to that for the AHI group (fold effect = 0.84/year; 95% CI = 0.38 to 1.82; P = 0.59) but not statistically significant (Fig. 4B). The fold effects on the proportion of infected T cells associated with each additional year on therapy in T_N, T_{CM}, T_{TM}, and T_{EM} cells derived from the CHI group were also similar to those from the AHI group. The PB-derived T_{SCM} cells were obtained from four participants in the CHI group. The reduction in the proportions of infected T_{SCM} cells during each additional year on therapy was not statistically significant (fold effect = 0.70/year; 95% CI = 0.15 to 3.37; P = 0.56). Of interest, T_{EM} cells derived from the CHI group had strong evidence for a stable proportion of HIV-1-infected T cells during each additional year on therapy, with a fold effect of 1.01 per year (95% CI = 0.87 to 1.17). However, CXCR5$^+$ CCR6$^-$ (X5$^+$ R6$^-$) memory CD4$^+$ T cells derived from the CHI group had a substantially lower proportion of infected T cells associated with each additional year of ART.

The LN data were derived from 4 participants from the AHI group, with samples obtained from 3.6 to 7.3 years of therapy (Tables 1 and 2). For 8 participants from the

TABLE 1 Participant demographics and clinical samples

Pt[a]	Pre-ART plasma sampling time point[b]	On-ART plasma sampling time point (duration)[c]	ART (yrs)[d]	CD4 Count (cells/µL)[e]	Viral load (HIV-1 RNA copies/mL)[f]	Therapeutic regimen[g]	Samples
AHI group							
2443	0 m	1.5 m / 2.5 m / 3.5 m	3.6	771	<40	DRV, RTV, TDF/FTC, MVC	PB, LN, Gut
2302	1 m	1 m / 2 m	4.6	696	<40	FPV, RTV, TDF/FTC	PB, Gut
2303	0 m	12 d / 1 m	4.6 / 4.8	606 / 606	<40 / <40	EFV/FTC/TDF	LN / PB
2277	0 m	5.8 y	5.8	574	<40	EFV/FTC/TDF	PB, Gut
2370	0 m	20 d	5.6 / 6.4	1165 / 895	<40 / <40	ABC/3TC, ATV	LN / PB, Gut
2243	0 m	7 d / 1 m	7.2	474	<40	TDF/FTC, NVP	PB
2268	0 m	7.3 y	7.3	913	<40	EFV/FTC/TDF	PB, LN, Gut
2278	0 m	11 d / 1 m	7.8	883	<40	TDF/FTC, NVP	PB
2259	0 m	9.2 y	8.3 / 9.2	745 / 683	<40 / <40	FPV, TDF/FTC, RTV	Gut / PB
2286	0 m	9.1 y	9.1	381	<40	EFV/FTC/TDF	PB, Gut
2275	1 m	13.4 y	13.3 / 13.4 / 15.3	1094 / 1094 / 1842	<40 / <40 / <40	TDF/FTC, NVP	Gut / PB / PB[h]
2115	N/A	N/A	17.3	659	<40	TDF/FTC, NVP	PB, Gut
CHI group							
2452	1.6 y / 1 m	2 m / 3 m	3 / 3.2	604 / 604	<40 / <40	MVC, RGV, ETR	LN, Gut / PB
1292	4.7 y / 2 m	1 m / 1.5 m	3.5	746	<40	EFV/FTC/TDF	PB, LN
2470	10 m / 0 m	20 d / 2 m	4.5	602	<40	ATV, RTV, TDF/FTC	PB, LN
2467	5 m / 2 m	1 m / 2 m	4.6 / 4.8	488 / 488	<40 / <40	EFV/FTC/TDF	LN / PB
2469	1.6 y / 0 m	20 d / 1 m	5.7	1089	<40	ABC/3TC, ATV	PB, LN
2254	3 y, 1 y	6 y	6	956	<40	EFV/FTC/TDF	PB, Gut
2241	1.6 y / 1 m	2 m / 3 m	6.3 / 6.4	569 / 569	<40 / <40	EFV/FTC/TDF	LN, Gut / PB
2445	1.5 y / 0 m	N/A	9.4	824	<40	EFV/FTC/TDF	PB, Gut, LN
2450	0 m	N/A	9.7	632	3418	TDF/FTC, NVP	PB
2472	N/A	14 d	10.8	555	<40	DRV, RGV, 3TC, RTV	PB, LN
2518	N/A	N/A	15.3 / 15.4	432 / 339	<40 / <40	TDF, AZT/3TC, NVP	PB / Gut
2046	N/A	N/A	16.3	564	<40	ECV, EFV/FTC/TDF	PB, Gut
2013	N/A	N/A	17.3 / 17.4	819 / 386	<40 / <40	ABC/3TC, RGV	PB / Gut
2026	N/A	N/A	17.7 / 17.8	476 / 383	<40 / <40	RTV, DRV, TDF, ABC/3TC	PB / Gut

[a]Pt, participant.

[b]Years (y) and months (m) before ART initiation. Only the CHI group was used for analysis. NA, not applicable.

[c]Years, months, or days (d) after ART initiation. Only the CHI group was used for analysis. NA, not applicable.

[d]Duration on ART at the time of sample isolation.

[e]CD4 cell count at the time of sample isolation.

[f]Viral load at the time of sample isolation.

[g]Therapeutic regimen at the time of sample isolation. 3TC, lamivudine; ABC, abacavir; ATV, atazanavir; AZT, zidovudine; DRV, darunavir; ECV, entecavir; EFV, efavirenz; ETR, etravirine; FPB, fosamprenavir; FTC, emtricitabine; MVC, maraviroc; NVP, nevirapine; RGV, raltegravir; RTV, ritonavir; TDF, tenofovir.

[h]Excluded from analyses of the fold effect of the proportion of cells infected for cross-sectional analysis.

CHI group, the LN samples were obtained from 3 to 10.8 years of therapy (Tables 1 and 2). For these LN samples, we found an overall lower proportion of HIV-1-infected T cells associated with each additional year on therapy, with a fold effect of 0.15/year and 0.74/year for the AHI and CHI groups, respectively (Fig. 4A and B). For the AHI group,

TABLE 2 Numbers of CD4$^+$ T cells used for analysis

Pt[a]	PB							LN				Gut			
	T_N	T_{SCM}	T_{CM}	T_{TM}	T_{EM}	$R6^{+a}$	$X5^{+a}$	T_N	T_{CM}	T_{TM}	T_{EM}	T_N	T_{CTM}	T_{EM}	CD4
AHI															
2243	3,014,850	NA	3,809,056	4,673,276	2,404,690	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2302	1,330,864	NA	797,531	296,884	458,564	NA	NA	NA	NA	NA	NA	17,358	182,989	97,127	NA
2303	330,864	NA	375,309	335,246	261,200	NA	NA	2,022,983	59,288	59,318	592,693	NA	NA	NA	NA
2277	1,200,000	NA	100,960	45,432	55,479	NA	NA	NA	NA	NA	NA	237	21,115	10,603	NA
2370	3,000,000	NA	3,600,000	6,000,000	944,444	NA	NA	1,674,001	1,040,000	2,196,000	388,000	3,816	81,863	76,234	NA
2443	975,309	NA	906,173	822,222	866,667	NA	NA	154,854	84,387	97,637	36,802	8,578	153,252	106,970	NA
2268	6,500,000	NA	6,000,000	5,400,000	1,500,000	NA	NA	676,184	687,452	1,388,581	282,959	2,094	40,141	30,937	NA
2278	1,476,049	NA	2,000,259	1,512,837	800,539	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2259	6,000,000	NA	2,330,864	59,483	332,905	NA	NA	NA	NA	NA	NA	18,670	44,651	35,267	NA
2286	5,212,158	NA	256,498	424,160	26,475	NA	NA	NA	NA	NA	NA	76	8,036	9,645	NA
2275	10,429,091	8,364	5,851,628	1,210,685	2,255,922	249,158	302,761	NA	NA	NA	NA	3,709	79,197	50,592	NA
2115	4,404,848	8,444	1,694,276	439,394	1,515,152	311,448	267,116	NA	NA	NA	NA	NA	NA	NA	18,182
CHI															
2452	108,642	NA	88,889	102,058	108,724	NA	NA	171,016	143,775	119,718	180,179	2,867	65,317	40,361	NA
1292	3,008,265	NA	92,347	118,552	454,644	NA	NA	4,192,593	18,999	32,118	26,919	NA	NA	NA	NA
2470	497,456	NA	227,579	503,719	227,322	NA	NA	47,438	70,074	50,643	18,696	NA	NA	NA	NA
2467	1,399,891	NA	797,687	478,109	399,501	NA	NA	402,207	72,717	96,955	13,476	NA	NA	NA	NA
2469	2,002,843	NA	237,283	139,048	356,717	NA	NA	316,545	48,391	66,250	10,844	NA	NA	NA	NA
2254	1,160,165	NA	42,798	52,675	70,584	NA	NA	NA	NA	NA	NA	62	11,274	11,499	NA
2241	59,259	NA	118,519	87,259	122,902	NA	NA	115,681	69,333	147,793	20,622	2,700	24,200	43,700	NA
2445	2,533,333	NA	111,111	83,951	50,405	NA	NA	2,463,674	479,499	353,437	98,905	544	10,998	51,545	NA
2450	1,133,333	NA	250,206	22,359	59,991	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2472	6,185,766	NA	731,875	771,811	1,393,073	NA	NA	361,722	264,048	656,120	136,488	NA	NA	NA	NA
2518	1,143,939	8,613	1,869,809	2,047,138	6,810,236	278,788	454,545	NA	NA	NA	NA	NA	NA	NA	18,182
2046	196,566	4,175	54,471	9,577	38,908	95,623	99,663	NA	NA	NA	NA	NA	NA	NA	18,182
2013	699,717	69,662	787,027	539,237	1,836,913	266,264	132,300	NA	NA	NA	NA	NA	NA	NA	153,535
2026	384,000	218,182	1,097,222	171,212	970,202	137,535	185,152	NA	NA	NA	NA	NA	NA	NA	90,909
Avg	2,384,711	52,907	1,267,755	975,790	900,821	223,136	240,256	1,049,908	253,164	438,714	150,549	5,059	60,253	47,040	59,798

[a]Pt, participant; R6$^+$, X5$^-$ R6$^+$ T cell subset; X5$^+$, X5$^+$ R6$^-$ T cell subset.

TABLE 3 Cellular markers used for a broad spectrum of CD4$^+$ T cell subsets sorted from PB, LNs, and gut

Anatomic site	CD4$^+$ T cell subset	Cell markers	Pt IDs^a
PB	T$_N$	CD3$^+$ CD4$^+$ CD45RO$^-$ CCR7$^+$ CD27$^+$ CD57$^-$	2243, 2302, 2303, 2277, 2370, 2443, 2268, 2278, 2259, 2286, 2452, 1292, 2470, 2467, 2469, 2254, 2241, 2445, 2450, 2472
		CD3$^+$ CD4$^+$ CD45RA$^+$ CD45RO$^-$ CCR7$^+$ CD27$^+$ CD127$^+$ CD95$^-$	2275, 2115, 2518, 2046, 2013, 2026
	T$_{SCM}$	CD3$^+$ CD4$^+$ CD45RA$^+$ CD45RO$^-$ CCR7$^+$ CD27$^+$ CD127$^+$ CD95$^+$	2275, 2115, 2518, 2046, 2013, 2026
	T$_{CM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CCR7$^+$ CD27$^+$	2243, 2302, 2303, 2277, 2370, 2443, 2268, 2278, 2259, 2286, 2452, 1292, 2470, 2467, 2469, 2254, 2241, 2445, 2450, 2472
		CD3$^+$ CD4$^+$ CD45RA$^-$ CD45RO$^+$ CCR7$^+$ CD27$^+$	2275, 2115, 2518, 2046, 2013, 2026
	T$_{TM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CCR7$^-$ CD27$^+$	2243, 2302, 2303, 2277, 2370, 2443, 2268, 2278, 2259, 2286, 2452, 1292, 2470, 2467, 2469, 2254, 2241, 2445, 2450, 2472
		CD3$^+$ CD4$^+$ CD45RA$^-$ CD45RO$^+$ CCR7$^-$ CD27$^+$	2275, 2115, 2518, 2046, 2013, 2026
	T$_{EM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CCR7$^-$ CD27$^-$	2243, 2302, 2303, 2277, 2370, 2443, 2268, 2278, 2259, 2286, 2452, 1292, 2470, 2467, 2469, 2254, 2241, 2445, 2450, 2472
		CD3$^+$ CD4$^+$ CD45RA$^-$ CD45RO$^+$ CCR7$^-$ CD27$^-$	2275, 2115, 2518, 2046, 2013, 2026
	X5$^+$ R6$^-$	CD3$^+$ CD4$^+$ CD45RA$^-$ CD45RO$^+$ CXCR5$^+$ CCR6$^-$	2275, 2115, 2518, 2046, 2013, 2026
	X5$^-$ R6$^+$	CD3$^+$ CD4$^+$ CD45RA$^-$ CD45RO$^+$ CXCR5$^-$ CCR6$^+$	2275, 2115, 2518, 2046, 2013, 2026
LNs	T$_N$	CD3$^+$ CD4$^+$ CD45RO$^-$ CCR7$^+$ CD27$^+$ CD57$^-$	2303, 2370, 2443, 2268, 2452, 1292, 2470, 2467, 2469, 2241, 2445, 2472
	T$_{CM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CCR7$^+$ CD27$^+$	2303, 2370, 2443, 2268, 2452, 1292, 2470, 2467, 2469, 2241, 2445, 2472
	T$_{TM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CCR7$^+$ CD27$^+$	2303, 2370, 2443, 2268, 2452, 1292, 2470, 2467, 2469, 2241, 2445, 2472
	T$_{EM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CCR7$^-$ CD27$^-$	2303, 2370, 2443, 2268, 2452, 1292, 2470, 2467, 2469, 2241, 2445, 2472
Gut	T$_N$	CD3$^+$ CD4$^+$ CD45RO$^-$ CD27$^+$	2302, 2277, 2370, 2443, 2268, 2259, 2286, 2275, 2452, 2254, 2241, 2445
	T$_{CTM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CD27$^+$	2302, 2277, 2370, 2443, 2268, 2259, 2286, 2275, 2452, 2254, 2241, 2445
	T$_{EM}$	CD3$^+$ CD4$^+$ CD45RO$^+$ CD27$^-$	2302, 2277, 2370, 2443, 2268, 2259, 2286, 2275, 2452, 2254, 2241, 2445
	CD4$^+$	CD3$^+$ CD4$^+$ CD8$^-$ CD14$^-$	2115, 2518, 2046, 2013, 2026

^{a}Identifiers (IDs) of participants included in the particular cell sorting.

the fold effect per year on therapy was influenced by all the T cell subsets sorted from LNs (Fig. 4A). For the CHI group, all the T cell subsets sorted from LNs also had a lower proportion of infected cells per year on ART, with a fold effect of 0.62 to 0.88 per year on therapy (Fig. 4B). Overall, our results provide some evidence that HIV-1-infected T cells decay substantially during each additional year of therapy within the LN-derived T cell subsets we sorted from 4 participants in the AHI group compared to the 8 participants from the CHI group.

We also found a strong correlation between the proportion of HIV-1-infected T cells sorted from LNs and PB within the acute/early participants for whom we had both LN and PB samples available (LN and PB correlations within T cell subsets were 0.98 to 0.99 [Table 4]). Among the chronic participants who had both samples available, we also found a strong correlation between LNs and PB within the T$_N$ and T$_{CM}$ cell subsets (0.93 and 0.96, respectively). We used mixed-effects models of both PB and LN data together to obtain extrapolated fold effects per year in LN-derived T cell subsets. These models exploit the PB-LN correlations to estimate what the effects would have been for all 12 AHI and all 14 CHI participants. These models estimated an overall 13% reduction in the proportion of infected T cells during each additional year on therapy in LNs of the AHI group, with a fold effect of 0.87/year (95% CI = 0.43 to 1.77) (Fig. 4A). This lower proportion of cells infected per year on ART was influenced by T$_N$ (fold effect = 0.57/ year) and T$_{CM}$ (fold effect = 0.67/year) cells. In the CHI group, the fold effects per year
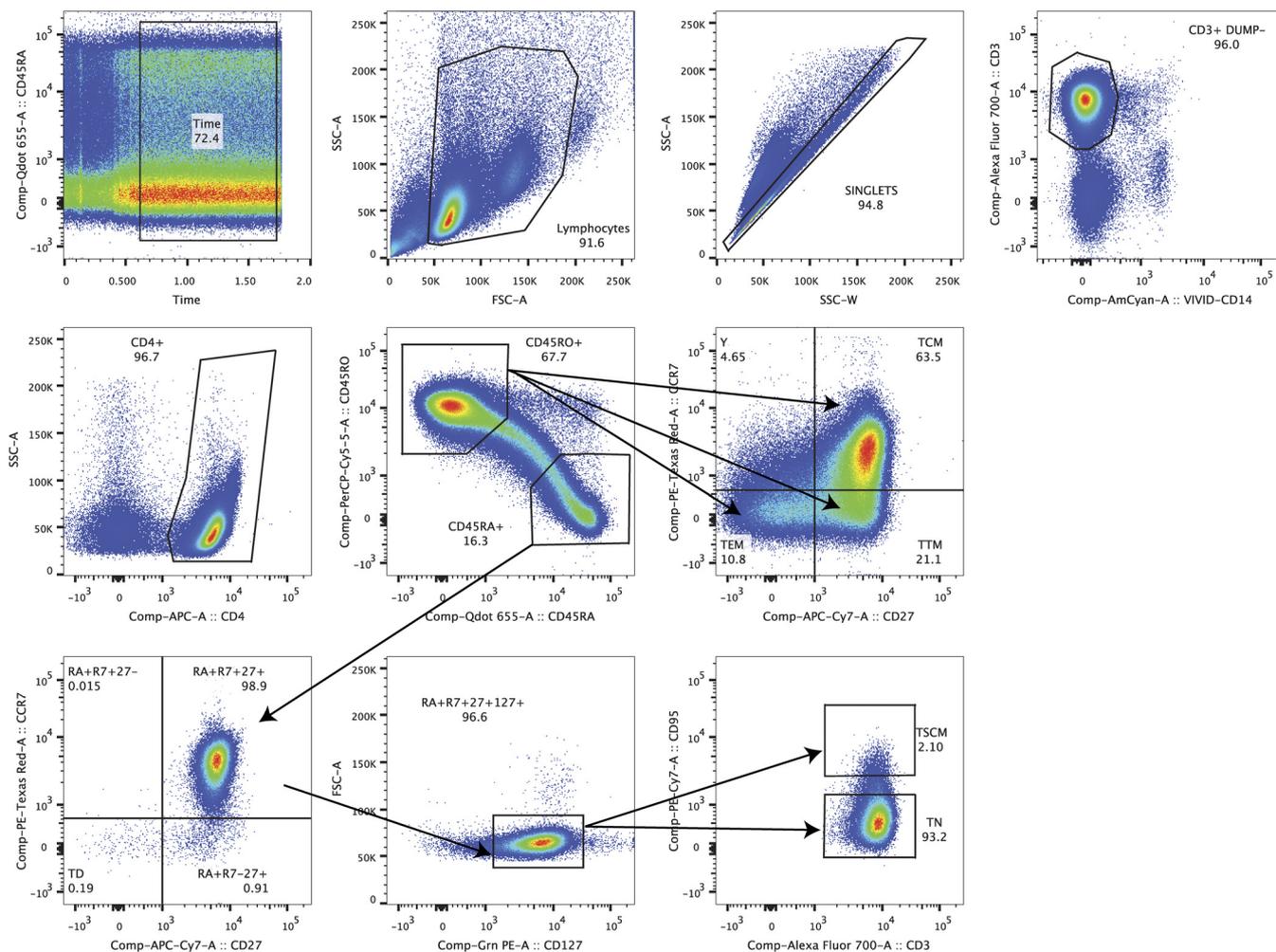
FIG 1 Example of the gating strategy for classical CD4$^+$ T cell populations from peripheral blood. Shown is the gating strategy for the CD4$^+$ T$_N$, T$_{SCM}$, T$_{CM}$, T$_{TM}$, and T$_{EM}$ cell subsets.

on therapy derived from the extrapolation of T cell subsets located in the LNs were comparable to those measured within the actual samples (0.86/year and 0.74/year, respectively) (Fig. 4B). This equates to a 14% reduction in the proportion of infected cells in the CHI group and was comparable to the overall reduction in LNs derived from the AHI group.

To assess the association between the ART duration and the proportion of infected cells within the gut tissue, we included nine and eight participants from the AHI and CHI groups, respectively (Tables 1 and 2). In the gut, the change in the proportion of HIV-1-infected T cells, as measured by fold effect, across all T cell subsets was similar to those for PB and LNs (AHI overall = 0.89/year; 95% CI = 0.64 to 1.25; P = 0.29; CHI overall = 0.65/year; 95% CI = 0.30 to 1.40; P = 0.17) but was not statistically significant for both the AHI and CHI groups (Fig. 4A and B). We also performed a correlation analysis within paired T cell subsets between gut and PB (Table 4). This correlation analysis was performed on participants who had more than 1,000 cells within each T cell subset derived from the gut (Table 2). In the AHI group, we found a moderate correlation with PB-derived T cell subsets within central/transitional memory T (T$_{CTM}$) and T$_{EM}$ cells sorted from the gut. For the CHI group, the correlations between gut and PB within T$_{CTM}$ and T$_{EM}$ cells were 0.79 and 0.93, respectively. However, only the T$_{EM}$ cell values were statistically significant.

To investigate the changes in cellular infection frequencies within each anatomic site in absolute terms, we determined the number of HIV-1-infected T cells per million
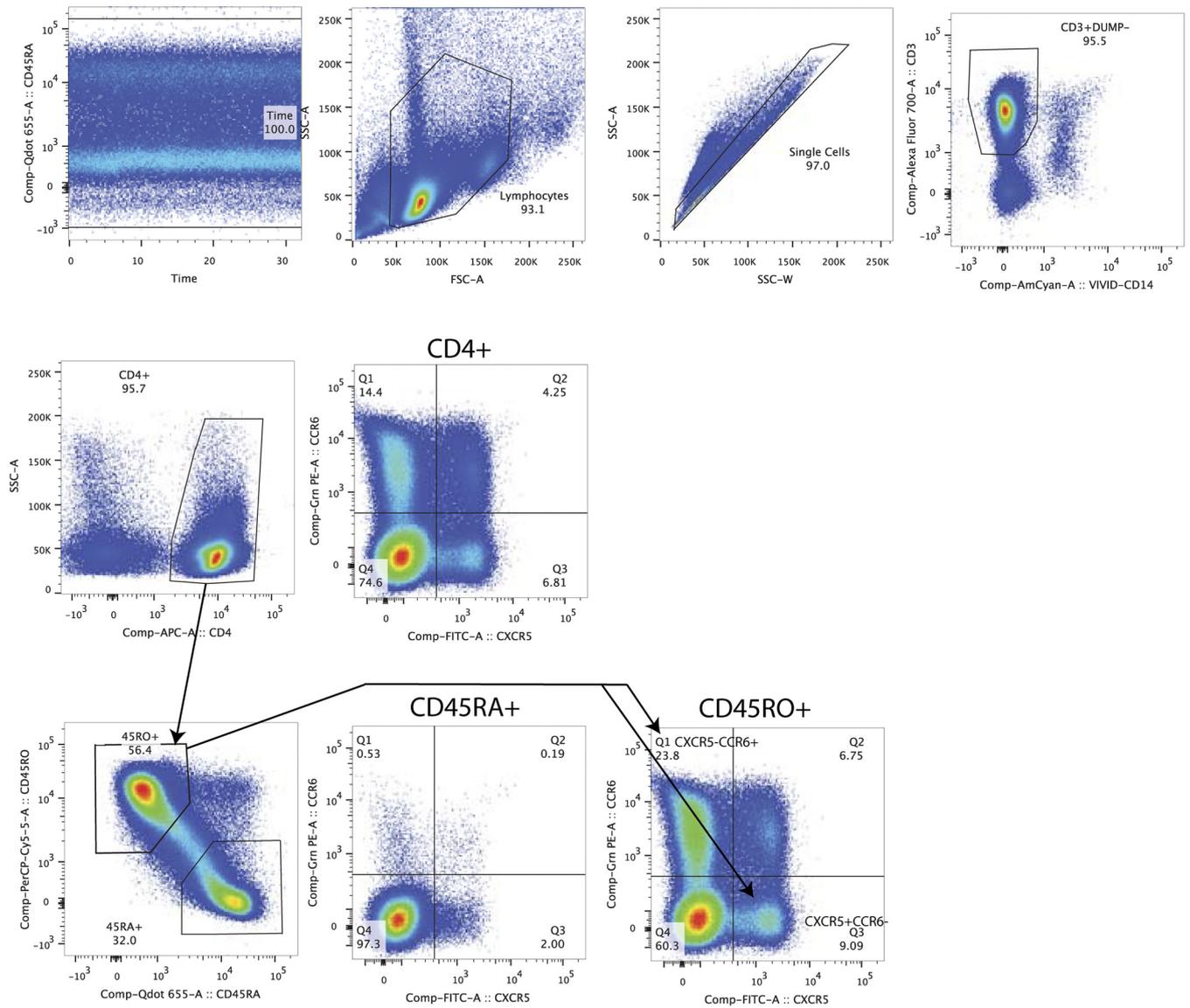
FIG 2 Example of the gating strategy for sorting CXCR5$^+$ CCR6$^-$ and CXCR5$^-$ CCR6$^+$ CD4$^+$ T cell memory populations from peripheral blood. Shown is the gating strategy for memory lymph node-homing (X5$^+$ R6$^-$) and memory gut-homing (X5$^-$ R6$^+$) CD4$^+$ T cell subsets.

within each T cell subset derived from PB, LNs, and gut for all participants (Fig. 5 and 6). These data recapitulated the results of the fold effect analysis per year on ART within each T cell subset.

Defective HIV-1 DNA in the p6-RT region during therapy. HIV-1 replication is characterized by rapid and highly error-prone reverse transcription, which lacks proof-reading capacity and causes genetic defects in viral genomes (15, 36–40). Also, an antiviral mechanism called G-to-A hypermutation renders in-frame stop codons and produces replication-deficient HIV-1 DNA during reverse transcription (15, 41, 42). These genetic defects have been identified in memory T cells during therapy (43–45). However, the impact of ART duration on the accumulation of genetically defective HIV-1 DNA is unclear. Therefore, we used a mixed-effects logistic regression model to calculate the proportion of genetically defective HIV-1 DNA p6-RT sequences during each additional year of therapy. Included in this analysis were 3,963 intracellular HIV-1 p6-RT sequences isolated from PB, LNs, and gut (Tables 5 and 6). HIV-1 sequences were defined as genetically defective due to insertions and/or deletions causing a frameshift, G-to-A hypermutations, stop codons, and/or internal deletions.
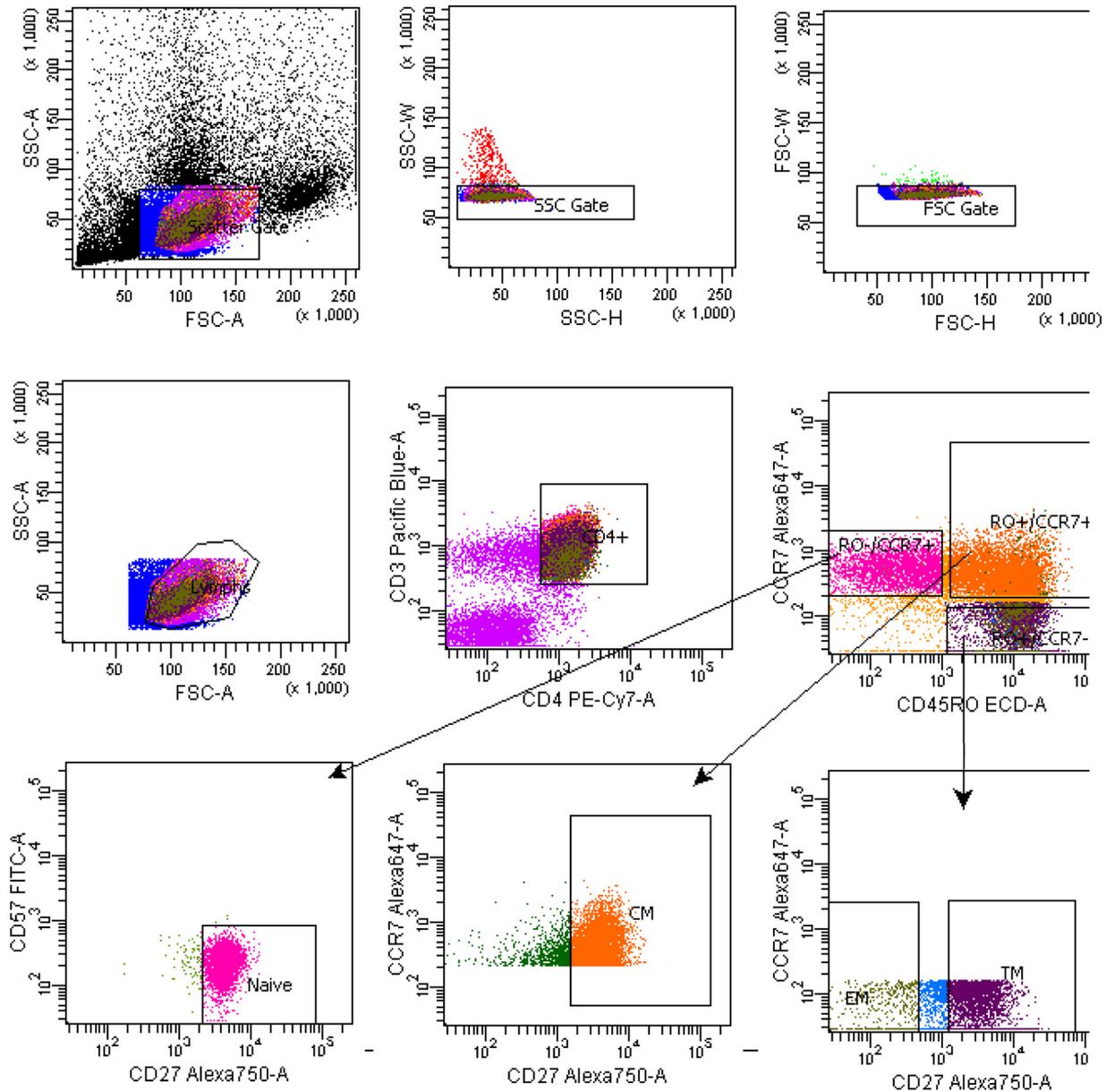
FIG 3 Example of the gating strategy for sorting classical CD4$^+$ T cell populations from lymph nodes. Shown is the gating strategy for CD4$^+$ T$_N$, T$_{CM}$, T$_{TM}$, and T$_{EM}$ cell subsets.

Overall, the odds of an HIV-1 DNA sequence being genetically defective did not substantially increase with the duration of ART within PB, LN, and gut tissues obtained from both participant groups (Fig. 7). Similar results were found in specific CD4$^+$ T cell subsets sorted from the anatomic sites obtained from the AHI group (Fig. 8) and the CHI group (Fig. 9). Our findings provide evidence that defective HIV-1 DNA p6-RT sequences do not increase substantially in the cell subsets and anatomic sites we analyzed during therapy regardless of when ART was initiated.

Pre- and on-therapy plasma samples contained few defective HIV-1 RNA p6-RT sequences. Studies have found that the majority of HIV-1 isolated from virions in the plasma is infectious and replication competent (46–48). Therefore, we compared the quantities of defective HIV-1 DNA sequences in T cell subsets sorted from PB, LNs, and gut sampled during ART to the 1,134 HIV-1 RNA sequences isolated from pre- and on-therapy plasma samples in the p6-RT region (Tables 5 and 6).

A

B
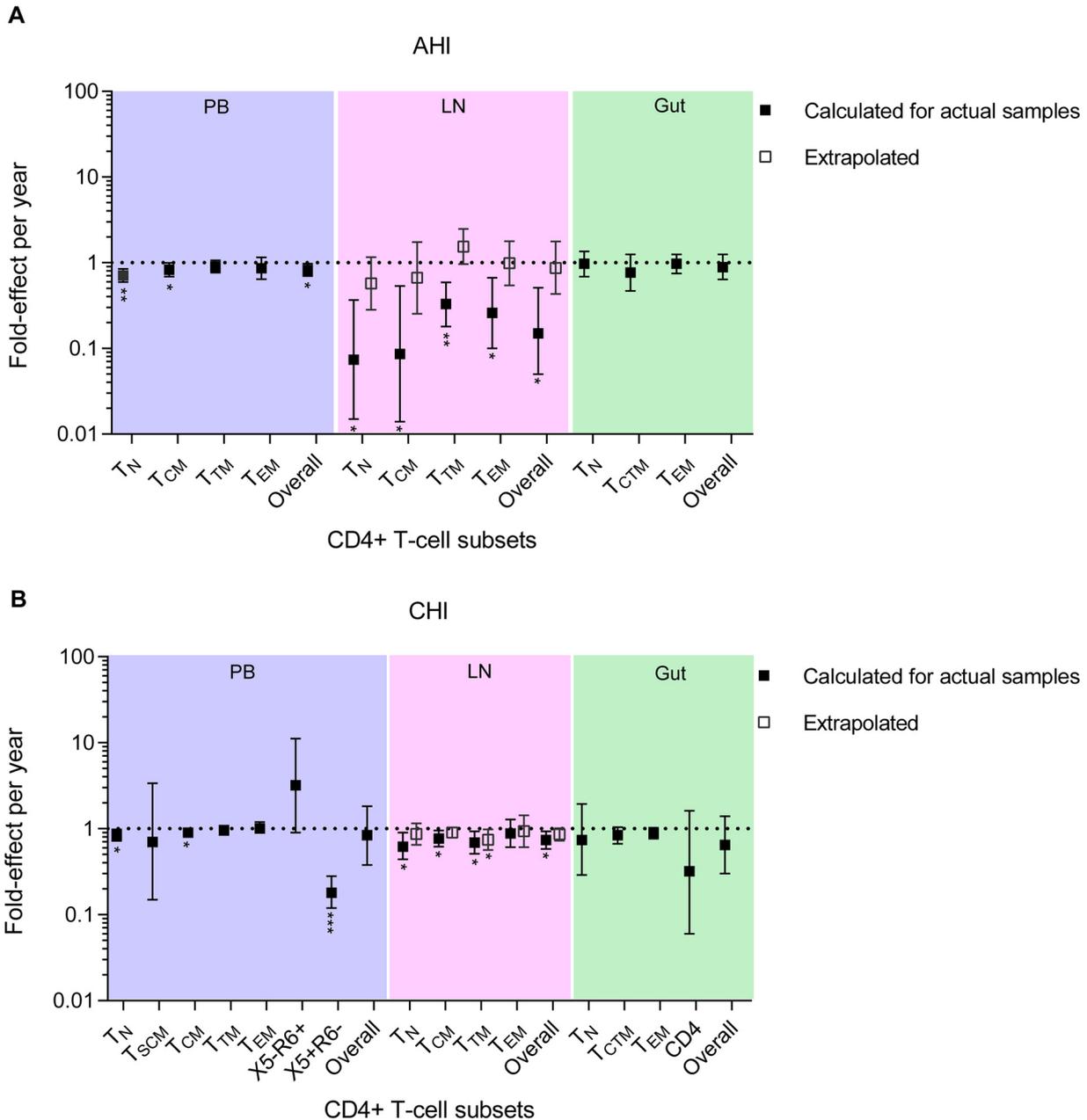
FIG 4 Fold effect of infection frequency during each year on ART. Shown is the fold effect per year of therapy on infection frequency in a broad range of CD4$^+$ T cell subsets sorted from PB (blue shading), LN (pink shading), and gut (green shading) tissues obtained from the AHI group (A) and the CHI group (B). The fold effect per year on therapy is a multiplicative effect that is equivalent to a fold change in the proportion of HIV-1-infected T cells from earlier to later time points during ART. For the LN, the fold effect per year on ART within 4 actual participant samples (with data from both PB and LNs) in the AHI group is indicated by the solid squares. The extrapolated fold effect per year on ART represents all 12 participants in the AHI group (open squares). For the LNs of the CHI group, the fold effect per year on ART derived from 8 actual participant samples (with data from both PB and LNs) is indicated by the solid squares, and the extrapolated fold effect per year of therapy for all 14 participants is indicated by the open squares. For each T cell subset, the fold effect per year on therapy was estimated when samples were available from at least 4 participants. The PB sample obtained from participant 2275 after 15.3 years of therapy was excluded from this cross-sectional analysis (Table 1). The error bars indicate the 95% confidence intervals of the infection frequency fold effect per year. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. The effects were estimated by negative binomial regression.

HIV-1 DNA sequences derived from the anatomic sites were more often genetically defective than HIV-1 RNA sequences derived from pretherapy plasma samples (Fig. 10). The odds that a PB-derived HIV-1 DNA sequence was genetically defective were about 3-fold higher than those for pretherapy plasma HIV-1 RNA sequences in both partici-

TABLE 4 Correlations between log proportions of HIV-1-infected T cells within PB, LNs, and gut

		95% CI		
Correlation	Estimate	Lower	Upper	P value
AHI group				
PB T_N with LN T_N	0.98	0.41	1.00	0.16
PB T_{CM} with LN T_{CM}	0.99	0.54	1.00	0.012
PB T_{TM} with LN T_{TM}	0.99	0.71	1.00	0.0068
PB T_{EM} with LN T_{EM}	0.99	0.50	1.00	0.012
PB T_N with gut T_N	−0.31	−0.86	0.58	0.95
PB T_{CM} and T_{TM} with gut T_{CTM}	0.75	0.10	0.95	0.032
PB T_{EM} with gut T_{EM}	0.62	−0.08	0.91	0.076
CHI group				
PB T_N with LN T_N	0.93	0.64	0.99	0.0009
PB T_{CM} with LN T_{CM}	0.96	0.78	0.99	0.0002
PB T_{TM} with LN T_{TM}	0.68	−0.05	0.94	0.065
PB T_{EM} with LN T_{EM}	0.34	−0.48	0.84	0.40
PB T_{CM} and T_{TM} with gut T_{CTM}	−0.45	−0.99	0.90	0.55
PB T_{EM} with gut T_{EM}	−0.46	−0.99	0.90	0.54

pant groups. In LNs, the odds that an HIV-1 DNA sequence was genetically defective were 2.5 and 3.0 times higher than those for pretherapy plasma sequences in the AHI and CHI groups, respectively. In the gut, the odds that a viral DNA sequence was defective were 4.9- and 2.1-fold higher than those for pretherapy plasma RNA se-
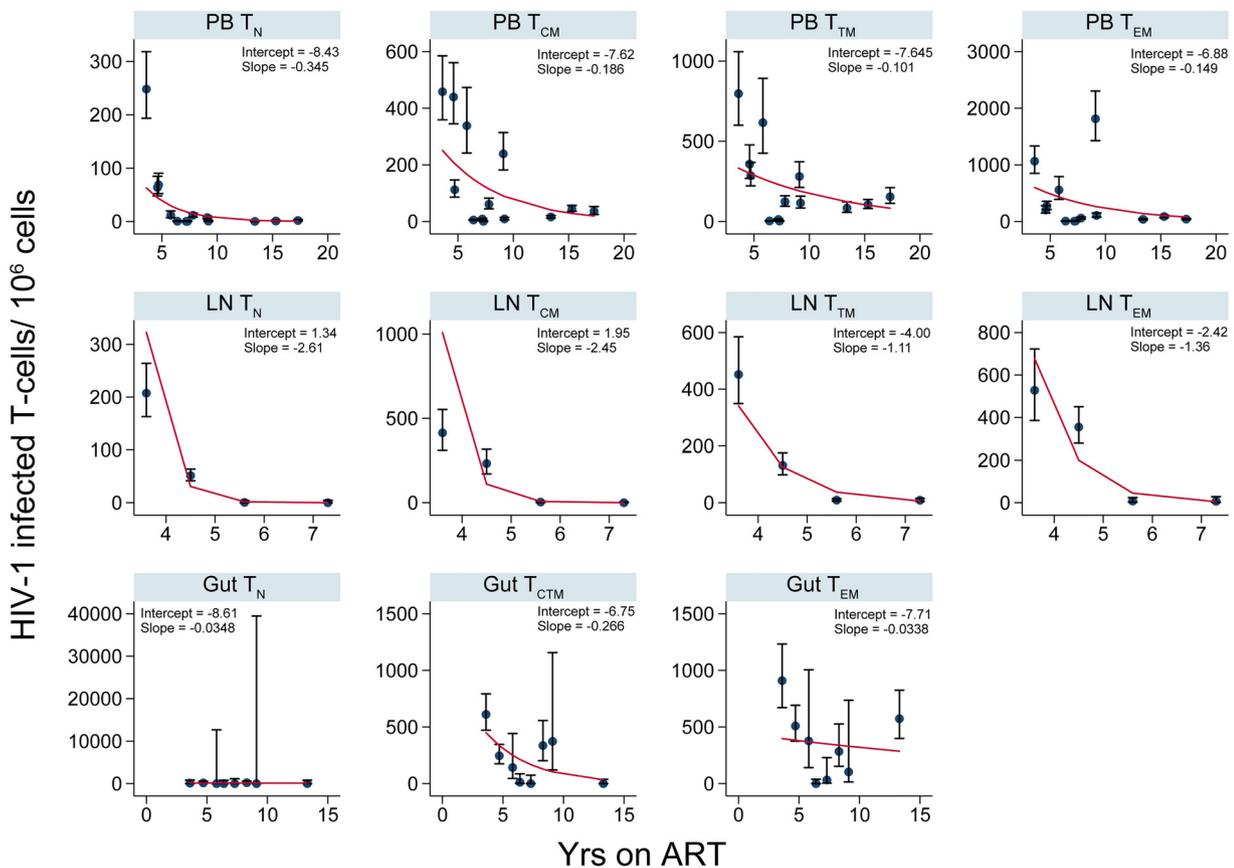
FIG 5 Numbers of HIV-1-infected T cells per million over years on ART (AHI group). The numbers of HIV-1-infected cells per million within CD4+ T cell subsets sorted from PB, LNs, and gut are shown. The error bars indicate the 95% confidence interval for each point. The red lines are derived from the fitted values calculated using the following equation at the sampling time points: $y = 1{,}000{,}000 \times \exp[\text{intercept} + \text{slope} \times (\text{years on ART})]$. The coefficients (intercept and slope) used to derive the fitted values for each T cell subset sorted from an anatomic site are shown.
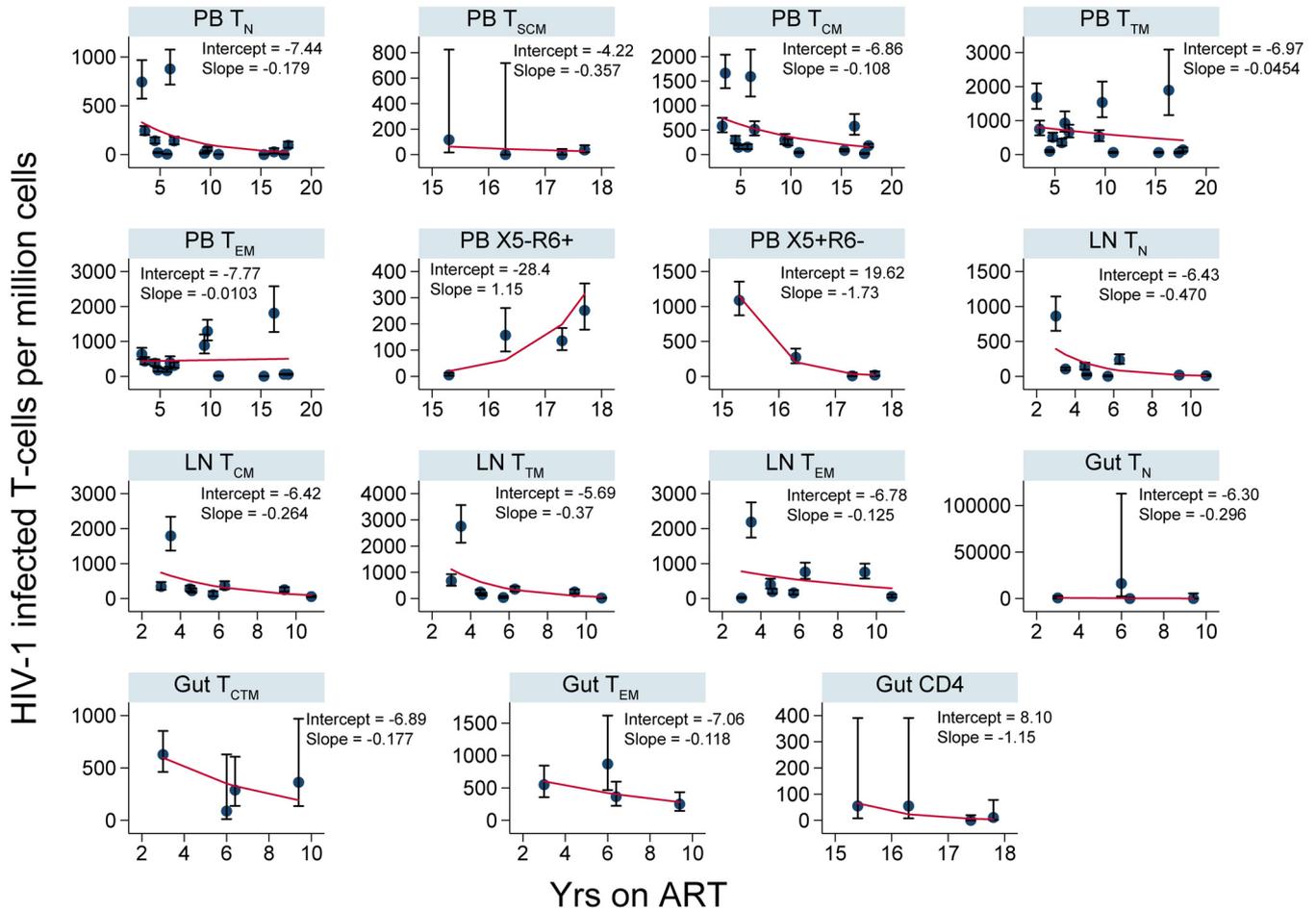
FIG 6 Numbers of HIV-1-infected T cells per million over years on ART (CHI group). The numbers of HIV-1-infected cells per million within CD4$^+$ T cell subsets sorted from PB, LNs, and gut are shown. The error bars indicate the 95% confidence interval for each point. The red lines are derived from the fitted values calculated using the following equation at the sampling time points: $y = 1{,}000{,}000 \times \exp[\text{intercept} + \text{slope} \times (\text{years on ART})]$. The coefficients (intercept and slope) used to derive the fitted values for each T cell subset sorted from an anatomic site are shown.

quences in the AHI and CHI groups, respectively. However, they did not reach statistical significance in the CHI group.

HIV-1 DNA sequences derived from the anatomic sites contained more genetic defects than viral RNA sequences derived from on-therapy plasma samples, but in some cases, they did not reach statistical significance (Fig. 10). In PB, the odds that a viral DNA sequence was defective were 5- to 6-fold greater than those for on-therapy plasma RNA sequences for both participant groups. Similarly, the odds that a viral DNA sequence was defective were about 5 times higher for LNs than for on-therapy plasma sequences in both participant groups. In the gut, the odds that a viral DNA sequence was genetically defective were 9.2-fold and 3.7-fold higher than for on-therapy plasma HIV-1 RNA sequences derived from the AHI and CHI groups, respectively. Furthermore, a majority of CD4$^+$ T cell subsets derived from the anatomic sites also showed higher odds that a viral sequence was genetically defective than pre- and on-therapy plasma-derived sequences in both the AHI and CHI groups (Fig. 11).

HIV-1 sequences from LN-derived T$_{EM}$ cells were more often genetically identical to pre- and on-therapy plasma viral RNA sequences. We found that HIV-1 RNA p6-RT sequences isolated from pre- and on-therapy plasma contained few genetic defects. The lack of defects within the p6-RT region of HIV-1 RNA sequences derived from pre- and on-therapy plasma indicated that these sequences came from replication-competent virions that could contribute to recrudescence. Therefore, we compared HIV-1 RNA and DNA p6-RT sequences to identify cell subsets that

TABLE 5 Numbers of HIV-1 sequences from plasma samples, CD4$^+$ T cell subsets, and anatomic sites (AHI group)

Pt[a]	Sequence type	Plasma[a]		PB[a]								LN					Gut				
		PT	OT	Total	T_N	T_{SCM}	T_{CM}	T_{TM}	T_{EM}	R6$^+$	X5$^+$	Total	T_N	T_{CM}	T_{TM}	T_{EM}	Total	T_N	T_{CTM}	T_{EM}	CD4
2243	Defective	5		19	0		9	9	1												
	Overall	37	0	83	1		30	35	17												
2302	Defective	1	0	30	3		9	5	13								14	0	6	8	
	Overall	35	0	112	31		26	28	27								54	3	24	27	
2303	Defective	3	0	21	2		9	8	2			20	2	4	5	9					
	Overall	29	12	95	27		25	21	22			123	31	30	30	32					
2277[b]	Defective	1	1	27	1		12	6	8								3	0	1	2	
	Overall	36	6	90	13		23	27	27								6	0	2	4	
2370	Defective	0	0	7	0		2	5	0			3	0	0	2	1	0	0	0	0	
	Overall	35	0	74	5		26	28	15			26	1	4	19	2	1	0	1	0	
2443	Defective	1	0	23	7		3	4	9			16	1	6	1	8	19	1	8	10	
	Overall	38	1	93	26		18	20	29			85	25	21	13	26	50	1	22	27	
2268[b]	Defective	2		3	0		0	2	1			2	0	0	2	0	0	0	0	0	
	Overall	36		25	1		2	10	12			10	0	0	9	1	1	0	0	1	
2278	Defective	5	0	15	3		3	5	4												
	Overall	33	0	90	24		23	24	19												
2259[b]	Defective	5	0	32	3		7	14	8								11	0	6	5	
	Overall	47	0	91	4		21	28	38								23	3	11	9	
2286[b]	Defective	7	0	16	2		7	5	2								0	0	0	0	
	Overall	48	1	105	23		26	26	30								3	0	2	1	
2275[b]	Defective	2	0	6	0	0	2	1	3	0	0						1	0	0	1	
	Overall	30	3	234	5	0	62	63	57	33	14						29	0	0	29	
2115	Defective			55	3	0	7	10	14	8	13						6				6
	Overall			124	8	0	18	23	25	22	28						7				7
Total	Defective	32	1	254	24	0	70	74	65	8	13	41	3	10	10	18	54	1	21	26	6
	Overall	404	23	1216	168	0	300	333	318	55	42	244	57	55	71	61	174	7	62	98	7

[a]Pt, participant; PT, pretherapy plasma; OT, on-therapy plasma; R6$^+$, X5$^-$ R6$^+$ T cell subset; X5$^+$, X5$^+$ 6$^-$ T cell subset.
[b]Source of HIV-1 RNA sequences (17, 23).

contained viral genomes most closely related to plasma-derived HIV-1. We compared the HIV-1 RNA and DNA sequences from the CHI group (Fig. 12), since HIV-1 sequences in participants treated during chronic infection are genetically diverse (17, 23, 24). A representative phylogenetic tree showing the genetic comparisons between HIV-1 RNA sequences derived from plasma and HIV-1 DNA sequences obtained during ART is presented in Fig. 13. We did not have pre- and on-therapy plasma samples collected from five of the CHI participants. Therefore, nine participants were included in the genetic comparison between HIV-1 DNA sequences isolated from the PB-derived T cell subsets and HIV-1 RNA sequences obtained from pretherapy or on-therapy plasma samples. For the comparison between the LN-derived T cell subsets and the plasma samples, LN tissue was collected from seven participants in the CHI group.

In PB, T_{CM} cells had the lowest odds that an HIV-1 DNA sequence would be genetically identical to a viral RNA sequence derived from pre- and on-therapy plasma samples compared to the other T cell subsets (Fig. 12A and C). In the LNs, T_{EM} cells had 4- and 5-times greater odds that an HIV-1 DNA sequence was genetically identical to viral sequences derived from pre- and on-therapy plasma samples, respectively, than T_N cells (Fig. 12B and D). We found that T_{EM} cells had higher odds that a viral sequence was genetically identical to pretherapy and on-therapy plasma HIV-1 RNA sequences than T_{CM} cells. Furthermore, we found 2.2-fold higher odds that a T_{EM}-derived HIV-1 DNA

TABLE 6 Numbers of HIV-1 sequences from plasma samples, CD4[+] T cell subsets, and anatomic sites (CHI group)

Pt[a]	Sequence type	Plasma[a]		PB[a]								LN					Gut				CD4
		PT	OT	Total	T_N	T_{SCM}	T_{CM}	T_{TM}	T_{EM}	$R6^+$	$X5^+$	Total	T_N	T_{CM}	T_{TM}	T_{EM}	Total	T_N	T_{CTM}	T_{EM}	
2452	Defective	3	0	19	6		4	2	7			11	3	5	3	0	6	1	2	3	
	Identical			45	2		0	22	21			9	0	0	9	0	16	0	8	8	
	Unique			74	28		26	7	13			77	29	29	16	3	34	2	23	9	
	Overall	65	7	125	30		29	30	36			91	29	30	27	5	52	2	33	17	
1292	Defective	3	1	18	3		6	5	4			22	5	8	8	1					
	Identical			14	4		2	0	8			26	6	2	5	13					
	Unique			85	27		26	17	15			91	28	27	22	14					
	Overall	58	30	110	34		31	21	24			125	36	30	29	30					
2470	Defective	9	0	14	4		7	1	2			22	3	8	10	1					
	Identical			25	5		4	3	13			21	4	2	9	6					
	Unique			46	8		15	11	12			47	11	14	15	7					
	Overall	71	3	85	18		21	19	27			86	20	22	26	18					
2467	Defective	1	3	27	6		8	6	7			19	4	9	3	3					
	Identical			19	4		2	0	13			13	4	3	2	4					
	Unique			73	16		24	18	15			41	1	20	12	8					
	Overall	70	36	111	23		29	27	32			67	9	30	14	14					
2469	Defective	4	0	19	3		1	3	12			9	1	0	1	7					
	Identical			30	0		3	8	19			5	0	2	0	3					
	Unique			59	14		21	10	35			19	1	2	3	13					
	Overall	63	7	96	14		28	19	35			29	1	6	3	19					
2254[b]	Defective	12	0	16	1		4	8	3								0	0	0	0	
	Identical			14	5		0	0	9								2	0	0	2	
	Unique			70	16		20	20	14								5	1	1	3	
	Overall	91	3	90	22		22	22	24								8	1	1	6	
2241	Defective	3	0	6	0		2	2	2			15	6	4	3	2	0	0	0	0	
	Identical			33	0		3	22	8			21	2	6	4	9	8	0	0	8	
	Unique			79	26		26	11	16			107	28	33	27	19	10	0	5	5	
	Overall	68	1	128	33		31	34	30			134	35	35	31	33	20	0	6	14	
2445[b]	Defective	10		32	6		6	6	14			15	3	3	4	5	5	0	1	4	
	Identical			30	7		0	6	17			14	0	0	2	12	0	0	0	2	
	Unique			64	20		22	16	6			82	23	25	23	11	10	0	3	7	
	Overall	77		106	27		26	26	27			108	27	27	26	28	14	0	4	10	
2450[b]	Defective	15		96	6		24	7	59												
	Identical			87	0		28	3	56												
	Unique			60	22		8	24	6												
	Overall	57		148	22		37	27	62												
2472	Defective		0	11	0		4	5	2			2	0	1	0	1					
	Identical			29	2		4	8	15			17	0	15	0	2					
	Unique			39	4		20	7	8			11	0	5	3	3					
	Overall		0	75	8		25	18	24			36	1	21	7	7					
2518	Defective			5	0	0	2	0	2	0	1						0				0
	Identical			46	0	0	4	11	19	3	9						0				0
	Unique			72	0	0	21	16	14	2	18						1				1
	Overall			131	0	1	28	28	36	5	33						1				1
2046	Defective			9	0	0	2	2	0	0	5						0				0
	Identical			37	3	0	0	5	22	7	0						0				0
	Unique			58	2	0	13	16	9	4	14						0				0
	Overall			122	5	0	28	25	25	13	26						0				0
2013	Defective			90	0	0	6	3	62	19	0						0				0
	Identical			80	0	0	5	11	50	11	3						0				0
	Unique			39	0	0	5	13	10	10	1						0				0
	Overall			123	0	0	10	26	62	21	4						0				0

(Continued on next page)

TABLE 6 (Continued)

Pt[a]	Sequence type	Plasma[a] PT	OT	PB[a] Total	T_N	T_{SCM}	T_{CM}	T_{TM}	T_{EM}	R6+	X5+	LN Total	T_N	T_{CM}	T_{TM}	T_{EM}	Gut Total	T_N	T_{CTM}	T_{EM}	CD4
2026	Defective			14	4	0	2	3	1	3	1						0				0
	Identical			29	3	2	0	5	22	7	0						0				0
	Unique			63	20	4	10	8	3	17	1						1				1
	Overall			107	23	6	13	13	25	25	2						1				1
Total[c]	Defective	60	4	376	39	0	78	53	177	22	7	115	25	38	32	20	11	1	3	7	0
	Identical			528	35	2	55	104	292	28	12	126	16	30	31	49	26	0	8	20	0
	Unique			881	203	4	257	194	176	33	34	475	121	155	121	78	61	3	32	24	2
	Overall	620	87	1,557	259	7	358	335	469	64	65	676	158	201	163	154	96	3	44	47	2

[a]Pt, participant; PT, pretherapy plasma; OT, on-therapy plasma; R6+, X5− R6+ T cell subset; X5+, X5+ R6− T cell subset.
[b]Source of HIV-1 RNA sequences (17, 23).
[c]Total number of HIV-1 sequences for all 14 participants from the CHI group.

sequence was genetically identical to viral sequences derived from on-therapy plasma than T_{TM} cells. Taken together, our genetic comparisons showed that T_{CM} cells within PB were least likely to contain HIV-1 DNA sequences that were genetically identical to HIV-1 RNA sequences derived from pre- and on-therapy plasma samples compared to
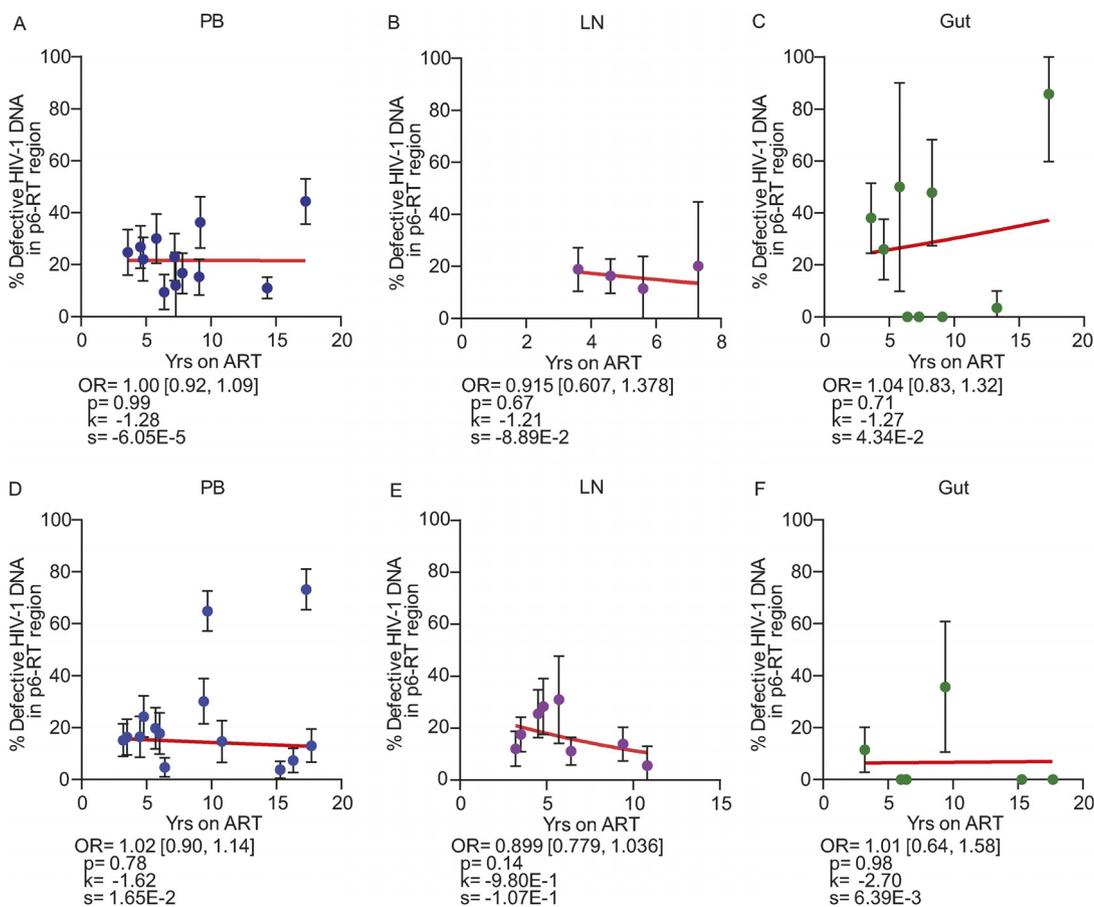
FIG 7 Proportions of defective HIV-1 DNA sequences in the p6-RT region during ART. Shown are the changes in the percentages of defective HIV-1 DNA p6-RT sequences during years on ART in PB (blue circles), LN (pink circles), and gut (green circles) tissues obtained from the AHI group (A to C) and the CHI group (D to F). The changes in the odds that a viral sequence is genetically defective during ART are indicated as OR per year on therapy. The 95% confidence intervals for the ORs are indicated in square brackets. Estimated by mixed-effects logistic regression, the fitted curves (red lines) follow the following equation: $y = 100 \times \{1 + \exp[-(k + xs)]\}^{-1}$, where y represents the proportions of defective HIV-1 DNA p6-RT sequences, k represents the y intercept, s represents the coefficient calculated as log(OR), and x represents the years on ART. Each HIV-1 DNA sequence is the unit of analysis, and the denominator for each sequence is a single HIV-1-infected T cell (33, 73). The confidence intervals of each data point were derived from the binomial distribution. The odds ratios and their confidence intervals were estimated by logistic regression models.
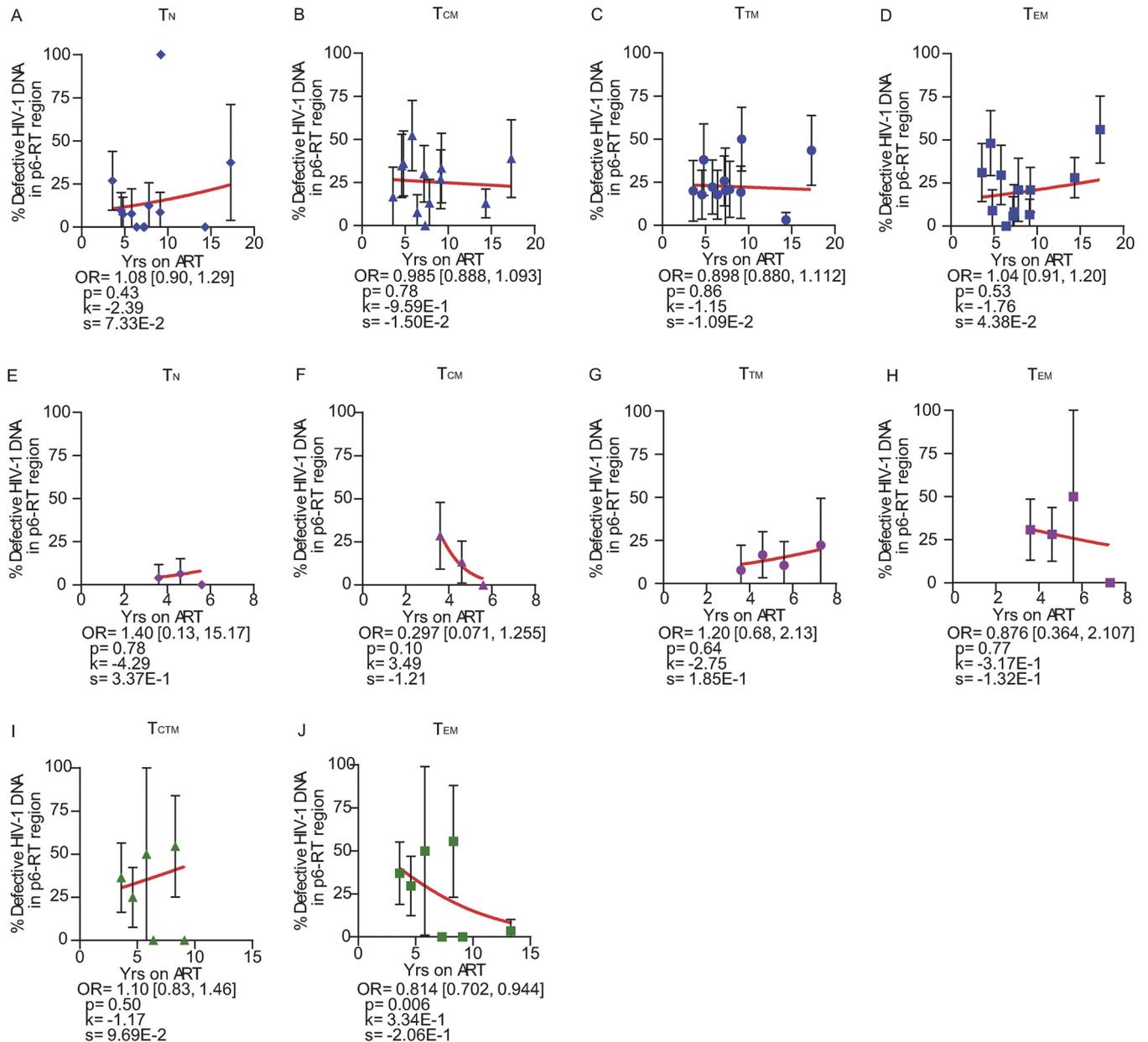
FIG 8 Proportions of defective HIV-1 DNA sequences from CD4+ T cell subsets during ART (AHI group). The effects of ART duration on the proportions of defective HIV-1 DNA p6-RT sequences in CD4+ T cell subsets sorted from PB (A to D), LNs (E to H), and gut (I to J) are indicated as OR per year on therapy. The 95% confidence intervals for the ORs are indicated in square brackets. Estimated by mixed-effects logistic regression, the fitted curves (red lines) follow the following equation: $y = 100 \times \{1 + \exp[-(k + xs)]\}^{-1}$, where y represents the proportions of defective HIV-1 DNA p6-RT sequences, k represents the y intercept, s represents the coefficient calculated as log(OR), and x represents the years on ART. Each HIV-1 DNA sequence is the unit of analysis, and the denominator for each sequence is a single HIV-1-infected T cell (33, 73). The confidence interval of each data point was derived from the binomial distribution. The odds ratios and their confidence intervals were derived by logistic regression models. Solid blue diamonds, T_N cells; solid blue triangles, T_{CM} cells; solid blue circles, T_{TM} cells; solid blue squares, T_{EM} cells, all derived from PB. Solid pink diamonds, T_N cells; solid pink triangles, T_{CM} cells; solid pink circles, T_{TM} cells; solid pink squares, T_{EM} cells, all derived from LNs. Solid green triangles, T_{CTM} cells; solid green squares, T_{EM} cells, all derived from the gut.

the other cell subsets. In contrast, LN-derived T_{EM} cells were most likely to contain HIV-1 DNA sequences that were genetically identical to HIV-1 RNA sequences derived from the plasma samples compared to other cell subsets.

Impact of ART duration on genetically identical and unique HIV-1 DNA sequences. To assess the contribution of cellular proliferation to viral persistence during ART, we used a mixed-effects logistic regression model to calculate the proportions of genetically identical and unique HIV-1 DNA p6-RT sequences with each additional year of ART for the CHI cohort. As these participants treated during chronic infection
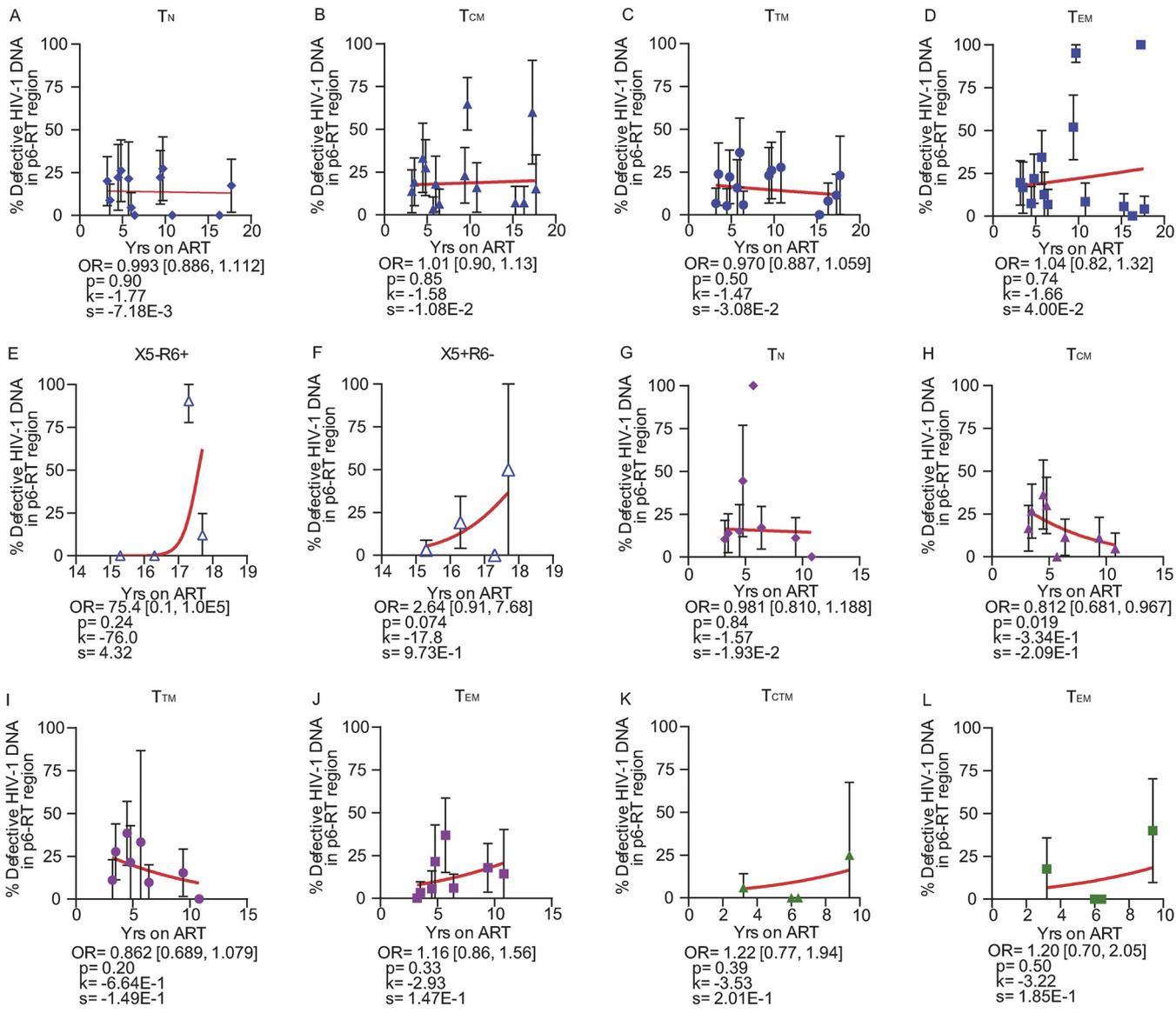
FIG 9 Proportions of defective HIV-1 DNA sequences from CD4+ T cell subsets during ART (CHI group). The effects of ART duration on the proportions of defective HIV-1 DNA p6-RT sequences in CD4+ T cell subsets sorted from PB (A to F), LNs (G to J), and gut (K to L) are indicated as OR per year on therapy. The 95% confidence intervals for the ORs are indicated in square brackets. Estimated by mixed-effects logistic regression, the fitted curves (red lines) follow the following equation: $y = 100 \times \{1 + \exp[-(k + xs)]\}^{-1}$, where y represents the proportions of defective HIV-1 DNA p6-RT sequences, k represents the y intercept, s represents the coefficient calculated as log(OR), and x represents the years on ART. Each HIV-1 DNA sequence is the unit of analysis, and the denominator for each sequence is a single HIV-1-infected T cell (33, 73). The confidence interval of each data point was derived from the binomial distribution. The odds ratios and their confidence intervals were estimated by logistic regression models. Solid blue diamonds, T_N cells; solid blue triangles, T_{CM} cells; solid blue circles, T_{TM} cells; solid blue squares, T_{EM} cells; open blue triangles, gut-homing (X5− R6+) cells; open blue triangles, lymph-homing (X5+ R6−) cells, all derived from PB. Solid pink diamonds, T_N cells; solid pink triangles, T_{CM} cells; solid pink circles, T_{TM} cells; solid pink squares, T_{EM} cells, all derived from LNs. Solid green triangles, T_{CTM} cells; solid green squares, T_{EM} cells, all derived from the gut.

contained genetically diverse HIV-1 (17, 23, 24), isolating genetically identical intracellular HIV-1 DNA sequences from these participants suggests the clonal expansion of HIV-1-infected cells. In this study, we define expansions of identical sequences (EIS) as ≥2 genetically identical sequences that are derived from the same cell subset sorted from an anatomic site to identify and quantify cellular proliferation within a specific T cell subset located in a particular anatomic site. Also, we defined unique HIV-1 DNA p6-RT sequences as those that are genetically distinct and exist on a single branch within a phylogenetic tree (Fig. 14). We included both genetically intact and defective HIV-1 DNA sequences in the p6-RT region in the analyses. We analyzed the impact of ART duration on the levels of identical HIV-1 DNA p6-RT sequences, which were derived
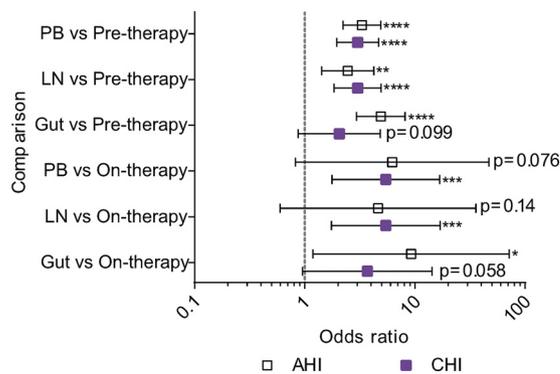
FIG 10 Odds that an HIV-1 sequence was defective in anatomic sites versus plasma samples. Shown is a comparison of the odds that an HIV-1 RNA sequence from pre- and on-ART plasma samples was defective versus the odds that a viral DNA sequence from PB, LN, and gut tissues was defective for the AHI group (open squares) and the CHI group (solid squares). The comparison of the odds is indicated as the OR; the error bars indicate the 95% confidence intervals for the ORs. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$. p6-RT sequences were used for the genetic comparisons. The odds ratios and their confidence intervals were estimated by logistic regression models.

from EIS, and unique viral sequences to further elucidate the dynamic interplay of HIV-1-infected cells in the viral reservoirs during therapy using mixed-effects logistic regression.

We observed that the proportions of identical HIV-1 DNA sequences increased while the proportions of unique viral sequences decreased with duration on therapy in PB and LNs (Fig. 15A, B, D, and E). In the gut, however, we found that the proportions of identical HIV-1 DNA sequences decreased while the proportions of unique viral sequences increased with duration on therapy when data from participants who were on ART for 3 to 17.8 years were included (Fig. 15C and F). However, the trends in LNs and gut did not reach statistical significance. Only PB showed a statistically significant increase in the odds that an HIV-1 DNA sequence was genetically identical to another by a factor of 1.09 per year on ART (95% CI = 1.03 to 1.15; $P = 0.003$) and a decrease in the odds that a viral sequence was unique by a factor of 0.941 (95% CI = 0.903 to 0.981; $P = 0.004$) (Fig. 15A and D). The phylogenetic analysis of HIV-1 DNA sequences derived from PB clearly showed increasing expansions of identical HIV-1 DNA sequences as ART duration increased (Fig. 15G to I).

We observed that genetically identical HIV-1 DNA sequences increased whereas unique viral sequences decreased in a majority of CD4$^+$ T cell subsets sorted from PB, LNs, and gut during therapy (Fig. 16 and 17). In PB-derived T$_{EM}$ cells, however, we observed the strongest evidence that the proportions of identical HIV-1 DNA sequences increased while the proportions of unique viral sequences decreased during each additional year on therapy (Fig. 16D). The odds of an HIV-1 DNA sequence being genetically identical to another in T$_{EM}$ cells increased with the duration of therapy by a factor of 1.11 per year (95% CI = 1.03 to 1.20; $P = 0.007$) (Fig. 16D). The odds that a viral sequence would be unique, however, decreased per year on ART by 0.909-fold in T$_{EM}$ cells (95% CI = 0.850 to 0.971; $P = 0.004$) (Fig. 16D).

DISCUSSION

Understanding the impact of ART duration on the dynamics and genetic composition of HIV-1-infected cells is critical for effective curative strategies. We therefore performed detailed genetic analyses of HIV-1 RNA sequences derived from pre- and on-therapy plasma samples and HIV-1 DNA sequences derived from a broad range of CD4$^+$ T cell subsets sorted from blood, lymph node, and gut tissues from participants who initiated treatment during acute/early and chronic infection. These studies allowed us to elucidate how the duration of therapy affects the number of HIV-1-infected T cells and the genetic composition of the proviruses they contain. In this cross-sectional analysis, we analyzed samples obtained after at least 3 years of effective therapy when
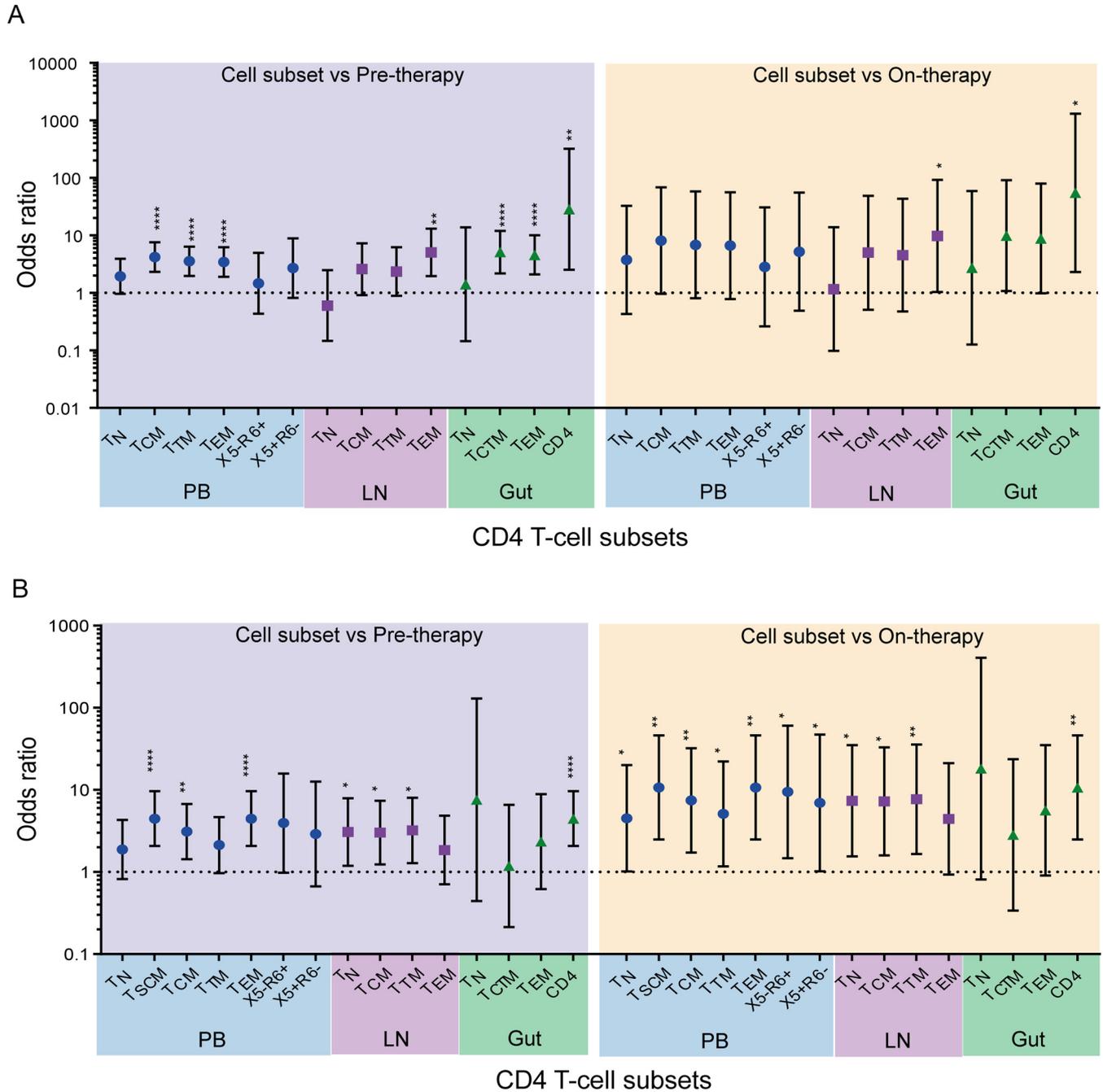
FIG 11 Odds that an HIV-1 sequence was defective in CD4$^+$ T cell subsets versus plasma samples. Shown are comparisons of the odds that a viral sequence was defective in CD4$^+$ T cell subsets sorted from PB, LN, and gut tissues derived from the AHI group (A) and the CHI group (B) to the odds that they were defective in pre- and on-therapy plasma samples. Blue shading, odds ratio of the CD4$^+$ T cell subset to pretherapy plasma; yellow shading, odds ratio of CD4$^+$ T cell subset to on-therapy plasma. *, $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$. p6-RT sequences were used for the comparisons. The odds ratios and their confidence intervals were derived by logistic regression models.

HIV-1 replication was fully suppressed to below the limit of detection to minimize the impact of episomal HIV-1 DNA in the single-proviral sequencing (SPS) data (49, 50).

We investigated the genetic composition of HIV-1 within the p6-RT region. It is well known that no subgenomic region can accurately represent the genetic diversity of full-length HIV-1 sequences (43, 44, 51). However, it has been shown that the *gag*-p6-*pro* region is better than other subgenomic regions at predicting the genetic diversity of full-length HIV-1 genomes in the viral populations derived from participants who initiated therapy during acute and chronic infection (51). The HIV-1 gene region we
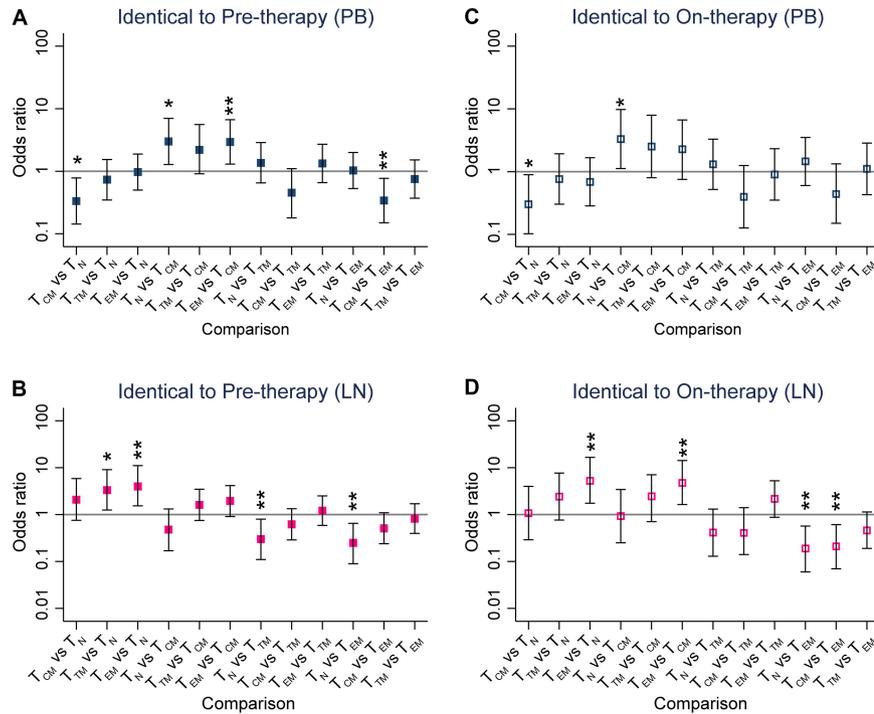
FIG 12 HIV-1 genetic comparisons between pre- and on-therapy plasma samples and T cell subsets from each anatomic site. Shown are comparisons of the odds that an HIV-1 DNA sequence was genetically identical to HIV-1 RNA sequences derived from pretherapy (A and B) or on-therapy (C and D) plasma samples. The odds ratios were derived from every pairwise comparison between CD4$^+$ T cell subsets sorted from PB (A and C) and LNs (B and D) obtained from the CHI group. The comparison of the odds is indicated as the OR, and the error bars indicate 95% confidence intervals of the ORs. *, $P < 0.05$; **, $P < 0.01$. p6-RT sequences were used for the genetic comparisons. The odds ratios and their confidence intervals were estimated by logistic regression models.

analyzed was longer than *gag*-p6-*pro* and had a clonal prediction score of 95, meaning that identical sequences in this region are identical throughout the viral genome with a probability of 95% (51). We analyzed over 5,000 HIV-1 RNA and DNA sequences, including both genetically intact and defective sequences, in this p6-RT region. Therefore, the subgenomic region selected for our genetic analyses, the clonal prediction score of the region, and the large number of sequences we analyzed increase the likelihood that the results of our study reflect the findings that would have resulted from full-length HIV sequence analyses. Moreover, the p6-RT HIV-1 DNA sequences that contain defects are not functional, as the region encodes viral enzymes important for replication. Previous studies have shown that defective HIV-1 genomes can produce viral proteins, including gag and pol (52, 53). Also, CD4$^+$ T cells from HIV-1-infected participants showed an increase in the expression of p24 when incubated with gag and pol viral proteins, indicating that these viral proteins can induce viral replication within infected cells (54). The p6-RT region we sequenced includes *gag* and *pol*, and based on the studies described above, we believe that quantifying how the proportions of defective or intact p6-RT sequences change over the duration of ART can help assess the risk of viral rebound and spread when ART is interrupted.

Several studies have demonstrated that the number of HIV-1-infected cells decays more rapidly in HIV-1-infected individuals who initiated therapy during early infection than in those who started treatment during chronic infection (55–57). Within the peripheral blood derived from the AHI group, we found an 18% reduction in the proportions of cells infected during each additional year of therapy. This result was similar to that of a recent longitudinal study that estimated the decay rate within CD4$^+$ T cells from peripheral blood at 13% per year on ART (55). However, one of the studies included participants who had several episodes of intermittent viremia per year,
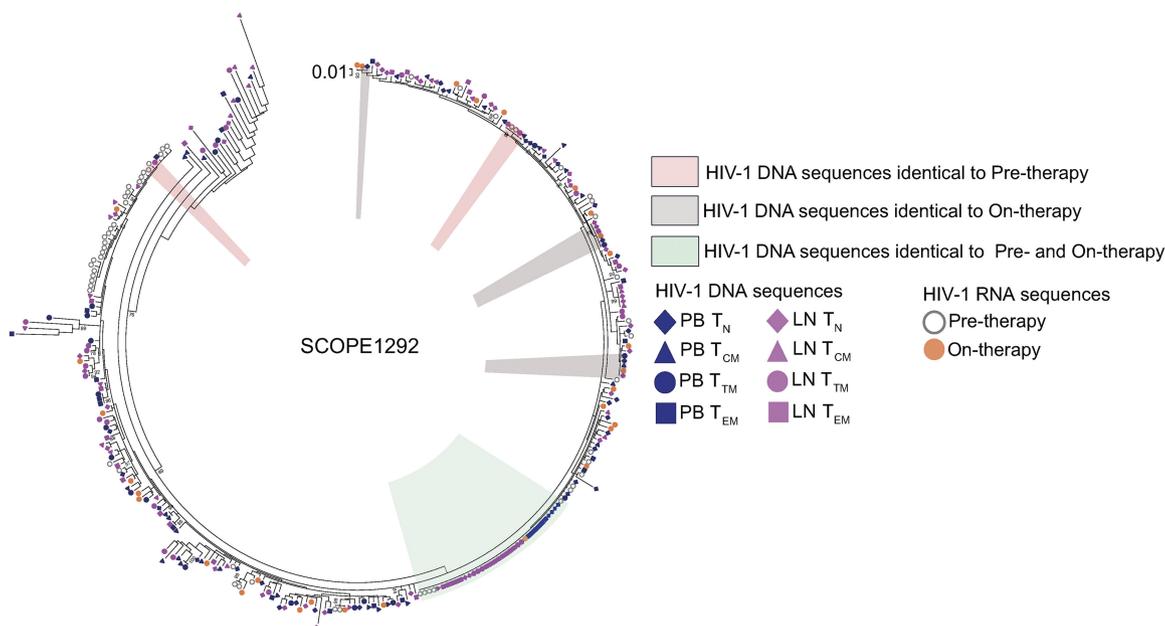
FIG 13 Representative phylogenetic tree showing HIV-1 sequence comparisons between plasma and CD4$^+$ T cell subsets. The phylogenetic regions containing HIV-1 DNA p6-RT sequences that were genetically identical to the viral sequences from the pretherapy plasma, to the on-therapy plasma, and to both of the plasma samples are marked. Individual HIV-1 DNA sequences derived from T_N (diamonds), T_{CM} (triangles), T_{TM} (circles), and T_{EM} (squares) cells sorted from PB (blue) and LNs (pink) are shown. Individual HIV-1 RNA sequences derived from pre- and on-therapy plasma are also shown. The maximum-likelihood phylogenetic tree was constructed with MEGA6.

indicating that their infections were not fully and continuously suppressed by the treatment (57). Moreover, the studies assessed the virus from total CD4$^+$ T cells (55, 57) or sorted a limited population of memory CD4$^+$ T cells (56) from peripheral blood when demonstrating a decay of HIV-1-infected cells during therapy. In contrast to these previous findings, we found similar decreases in the number of HIV-1-infected cells within the peripheral blood from both participant cohorts irrespective of whether treatment was initiated during acute/early or chronic infection. In both participant cohorts, this decrease in HIV-1-infected cells resulted from a statistically significant association of lower HIV-1 infection frequencies of naive and central memory T cells with each additional year of ART. However, in the case of participants treated during chronic infection, we found a large decrease in the proportion of HIV-1-infected cells over the years of therapy in the CXCR5$^+$ CCR6$^-$ cell subset. A recent study describes cells expressing PD-1 and CXCR5 as circulating follicular helper T cells (58). We did not sort for the PD-1 cell marker, but our data may suggest that HIV-1 infection of circulating follicular helper T cells is less stable and even decays over time. Moreover, it has been found that CXCR5$^+$ memory CD4$^+$ T cells sorted from peripheral blood coexpress CCR7, indicating these cells comprise a subpopulation of peripheral-blood-derived T_{CM} cells (59). Thus, the CXCR5$^+$ CCR6$^-$ cell subset could also have contributed to the decrease in the number of HIV-1-infected T_{CM} cells during each year on therapy in the peripheral blood of the participants treated during chronic infection. Overall, these findings indicate that the numbers of HIV-1-infected naive and T_{CM} cells within the peripheral blood decrease during effective therapy. This observed decrease could be due to the decay of T cells that harbor defective proviruses, as cells contributing to the HIV-1 reservoir are known to contain predominantly defective viral genomes (44).

For AHI and CHI participants with PB and LN samples available for analysis, HIV-1-infected cells from the lymph node decreased per year on ART. A substantial decrease in the proportion of HIV-1-infected T cells during each additional year on ART was found for all lymph node-derived CD4$^+$ T cell subsets from four participants treated during acute/early infection with paired PB and LN samples. Extrapolation to all 12 acute/early participants estimated smaller reductions in the proportion of HIV-1-
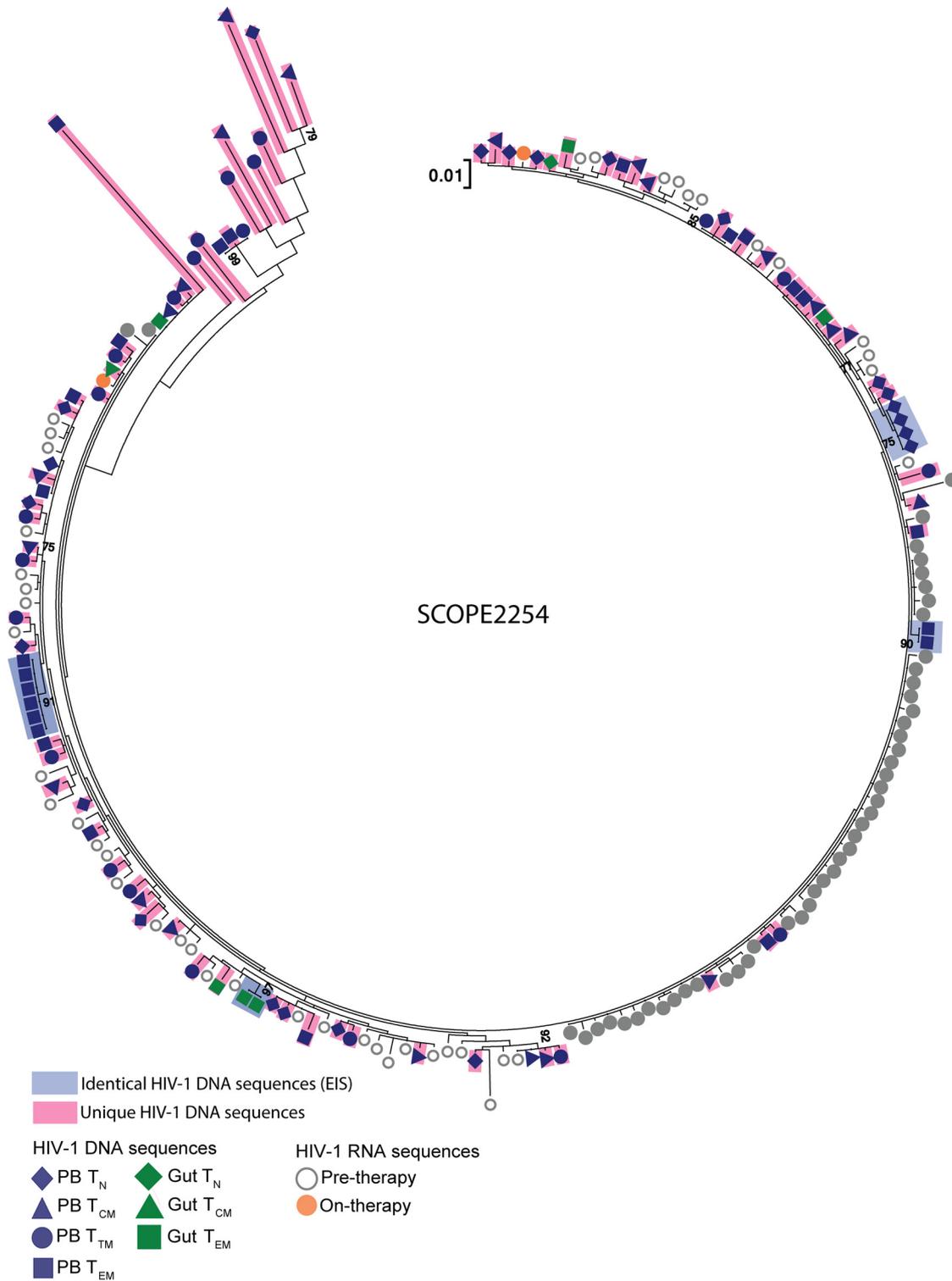
FIG 14 Representative phylogenetic tree showing identical and unique HIV-1 DNA sequences. Individual HIV-1 DNA p6-RT sequences that were genetically unique and located on single branches within a phylogenetic tree are colored pink. The unique viral sequences were genetically distinct and not identical to any other viral RNA and/or DNA sequence. Identical HIV-1 DNA sequences from EIS are colored blue. Individual HIV-1 DNA sequences derived from T_N (diamonds), T_{CM} (triangles), T_{TM} (circles), and T_{EM} (squares) cells sorted from PB (blue) are shown. Individual viral DNA sequences derived from T_N, T_{CTM}, and T_{EM} cells sorted from gut (green) are also shown. Individual HIV-1 RNA sequences derived from pre- and on-therapy plasma are shown. The maximum-likelihood phylogenetic tree was constructed with MEGA6.
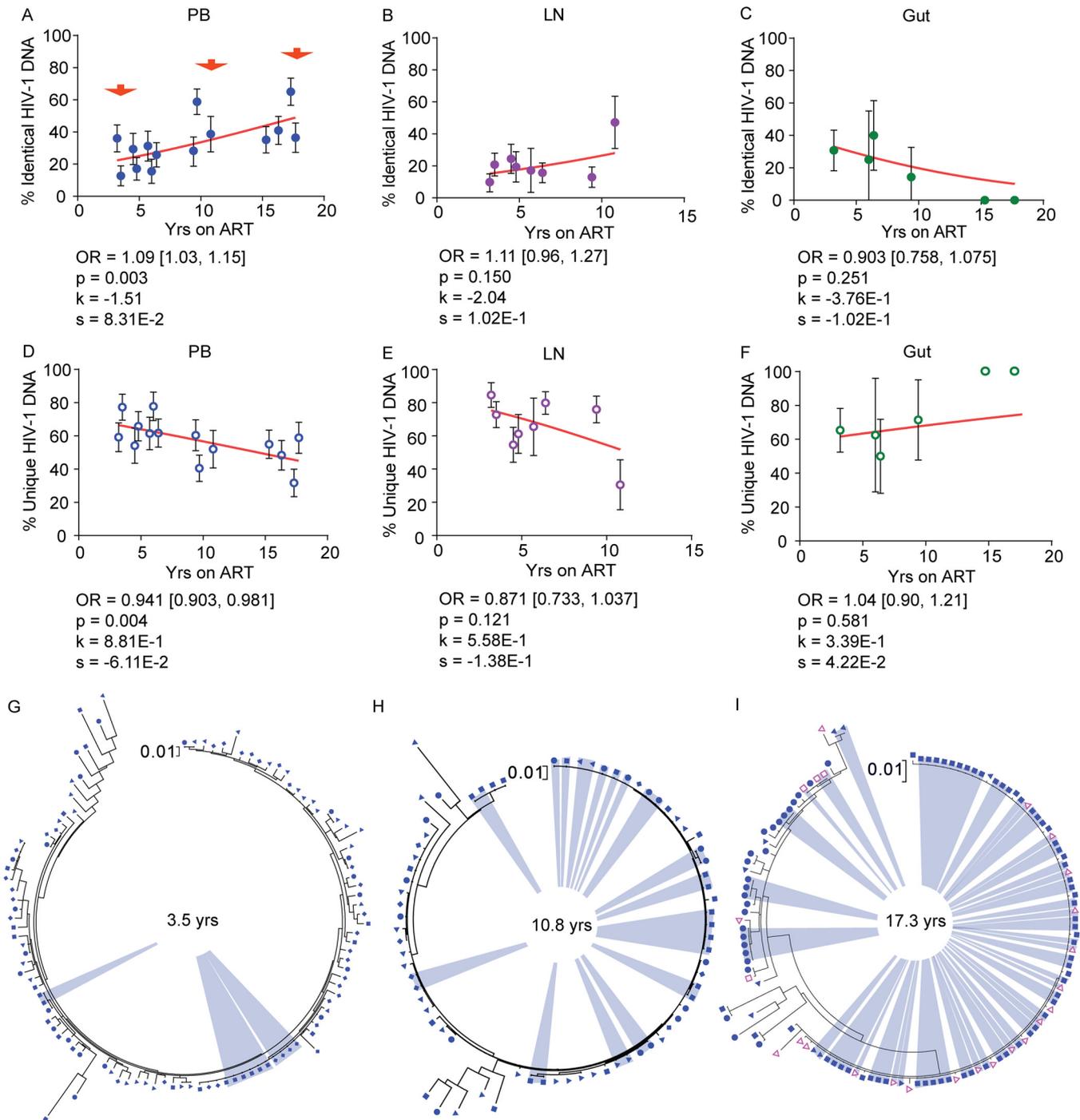
FIG 15 Genetically identical and unique HIV-1 DNA sequences derived from anatomic sites during ART. (A to F) The associations of the percentages of identical (A to C) and unique (D to F) HIV-1 DNA p6-RT sequences with years on ART in PB (A and D), LN (B and E), and gut (C and F) tissues derived from the CHI group are shown. The effects of therapy duration on the odds that a viral sequence was genetically identical are indicated as OR per year on therapy. The 95% confidence intervals for ORs are indicated in square brackets. Estimated by mixed-effects logistics regression, the fitted curves (red lines) follow the following equation: $y = 100 \times \{1 + \exp[-(k + xs)]\}^{-1}$, where y represents the proportions of genetically identical HIV-1 DNA p6-RT sequences, k represents the y intercept, s represents the coefficient calculated as log(OR), and x represents the years on ART. Each HIV-1 DNA sequence is the unit of analysis, and the denominator for each sequence is a single HIV-1-infected T cell (33, 73). The confidence interval of each data point was estimated from the binomial distribution. The confidence interval of each odds ratio was estimated by logistic regression models. (G to I) Maximum-likelihood phylogenetic trees indicating an increase in expansions of genetically identical HIV-1 DNA p6-RT sequences (blue) at 3.5 years (participant 1292) (G), 10.8 years (participant 2472) (H), and 17.3 years (participant 2013) (I) after ART initiation, as indicated in panel A (red arrows). The trees included HIV-1 DNA p6-RT sequences derived from T_N (diamonds), T_{CM} (solid triangles), T_{TM} (circles), T_{EM} (solid squares), $X5^+ R6^-$ (open squares), and $X5^- R6^+$ (open triangles) cells sorted from PB.
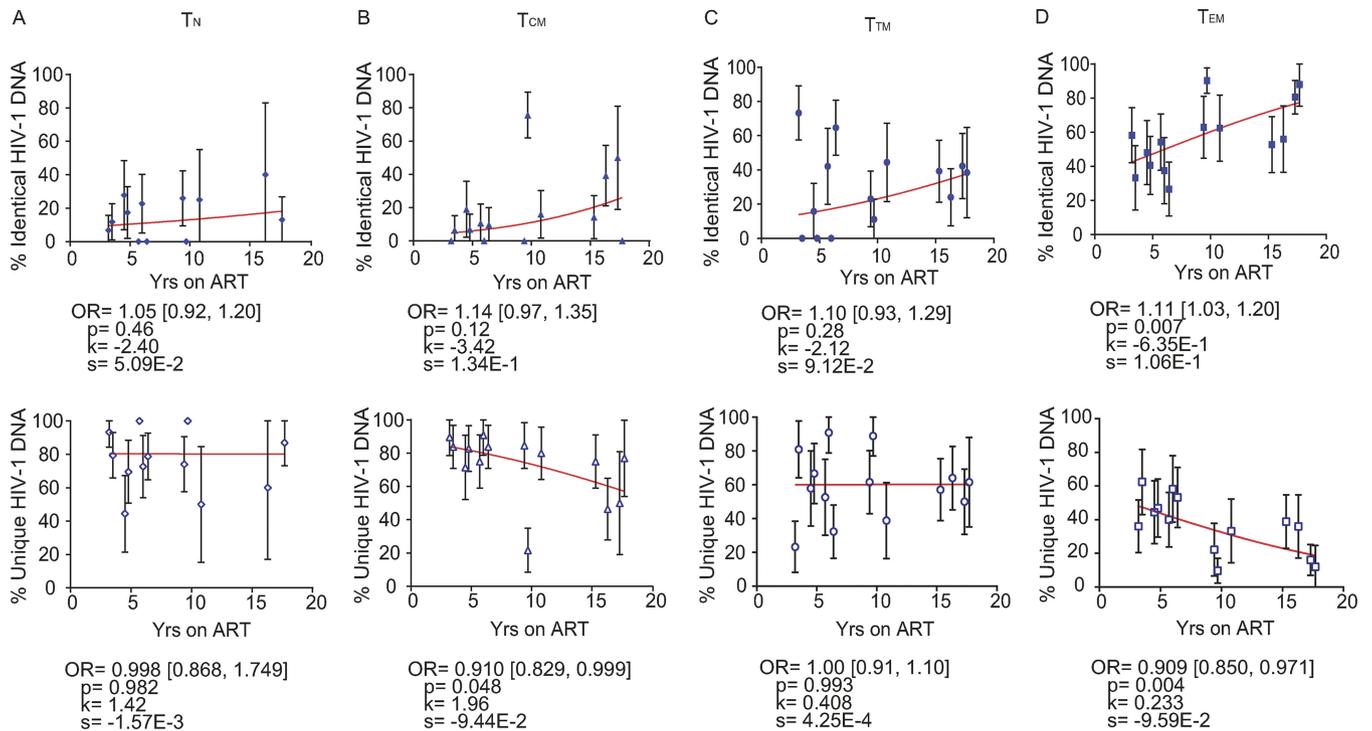
FIG 16 Genetically identical and unique HIV-1 DNA sequences from PB-derived CD4$^+$ T cell subsets during ART. The effects of ART duration on the percentages of identical (top) and unique (bottom) HIV-1 DNA p6-RT sequences derived from T_N (A), T_{CM} (B), T_{TM} (C), and T_{EM} (D) cells sorted from PB of the CHI group are indicated as OR per year on therapy. The 95% confidence intervals for the ORs are indicated in square brackets. Estimated by mixed-effects logistic regression, the fitted curves (red lines) follow the following equation: $y = 100 \times \{1 + \exp[-(k + xs)]\}^{-1}$, where y represents the proportions of genetically identical or unique HIV-1 DNA p6-RT sequences, k represents the y intercept, s represents the coefficient calculated as log(OR), and x represents the years on ART. Each HIV-1 DNA sequence is the unit of analysis, and the denominator for each sequence is a single HIV-1-infected T cell (33, 73) The confidence interval of each data point was estimated from the binomial distribution. The odds ratios and their confidence intervals were estimated by logistic regression models.

infected cells within LNs over 3 to 17.8 years of therapy that were similar to those for the participants treated during chronic infection. This finding suggests that tissue restoration and T cell reconstitution in lymph nodes reduced the number of HIV-1 cells during therapy for both participant groups (60).

Due to errors made during HIV-1 replication, genetic mutations that reduce viral fitness accumulate (15, 37, 38, 61). Genetic defects, such as G-to-A hypermutation, internal deletions, and nucleotide insertions and/or deletions that cause a frameshift, also accumulate in proviruses during viral replication (15, 37–42, 45, 62). However, our previous studies have shown a lack of HIV-1 evolution during ART, supporting a small amount of evidence for ongoing viral replication during therapy (17, 23). We combined all the defective features within the p6-RT region and found that the proportion of genetically defective HIV-1 DNA sequences did not appear to increase substantially with the duration of therapy. This was observed in most of the cell subsets and the anatomic sites we analyzed regardless of when ART was initiated. Our results revealing that defective HIV-1 sequences do not accumulate during therapy agree with those of a recent study showing that cytotoxic T cells can target cells containing defective proviruses (63). Taken together, our findings provide evidence that defective viral sequences are established during multiple rounds of viral replication before viral suppression rather than during effective therapy.

We found that plasma samples obtained before ART and during the early phase of treatment contained a lower proportion of defective viral sequences than HIV-1 DNA sequences derived from blood or tissue CD4$^+$ T cells. This result indicates that plasma-derived viral sequences most likely represent a population of HIV-1 that is infectious and capable of producing new virions (46–48). Compared to intracellular HIV-1 sequences from the other T cell subsets, T_{CM} cells from the peripheral blood were least
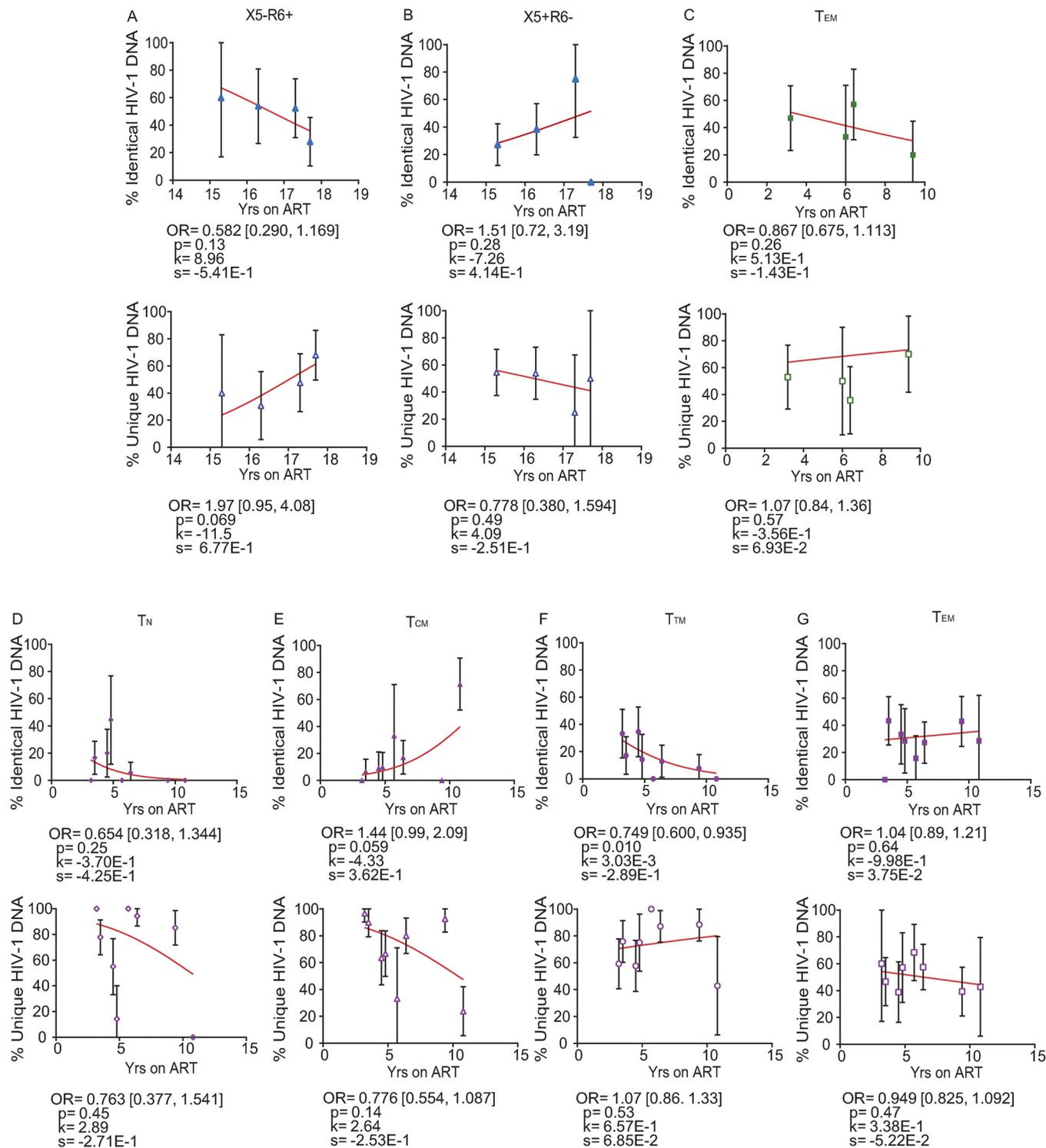
FIG 17 Genetically identical and unique HIV-1 DNA sequences from tissue-homing and tissue-derived CD4+ T cell subsets during ART. The effects of ART duration on the percentages of identical (top) and unique (bottom) HIV-1 DNA p6-RT sequences derived from CD4+ T cell subsets sorted from PB, LN, and gut tissues obtained from the CHI group are indicated as OR per year on therapy. The effects of ART duration on the proportions are shown in X5− R6+ (A) and X5+ R6− (B) cells sorted from PB; in T_{EM} cells (C) sorted from the gut; and in T_N (D), T_{CM} (E), T_{TM} (F), and T_{EM} (G) cells sorted from LNs. The 95% confidence intervals for the ORs are indicated in square brackets. Estimated by mixed-effects logistic regression, the fitted curves (red lines) follow the following equation: $y = 100 \times \{1 + \exp[-(k + xs)]\}^{-1}$, where y represents the proportions of genetically identical or unique HIV-1 DNA p6-RT sequences, k represents the y intercept, s represents the coefficient calculated as log(OR), and x represents the years on ART. Each HIV-1 DNA sequence is the unit of analysis, and the denominator for each sequence is a single HIV-1-infected T cell (33, 73). The confidence interval of each data point was derived from the binomial distribution. Odds ratios and their confidence intervals were estimated by logistic regression models.

likely to contain HIV-1 DNA sequences that were identical to pre- and on-therapy plasma RNA sequences. Our recent study revealed that T_{CM} cells contain the smallest amount of genetically intact HIV-1 proviruses in peripheral blood (43). In agreement with the previous study, our genetic analysis involving intracellular HIV-1 DNA sequences and plasma-derived viral RNA sequences suggests that the T_{CM} cell subset within the peripheral blood is the least likely cellular source for infectious HIV-1. Importantly, our genetic analyses revealed that, in the lymph node tissue, T_{EM} cells were highly enriched with viral sequences that were genetically identical to plasma-derived HIV-1 sequences. These results suggest that T_{EM} cells are a probable source for infectious HIV-1 in the lymph node compared to all other T cell subsets.

The presence of genetically identical HIV-1 DNA sequences indicates that cellular proliferation plays a role in the maintenance of persistent HIV-1 during therapy (17, 23, 25–28, 43, 64). Our in-depth genetic analysis revealed that the genetically identical HIV-1 DNA sequences increased whereas the unique viral sequences decreased in a majority of CD4$^+$ T cell subsets sorted from PB and LNs during therapy. This indicates that cellular proliferation compensates for the decay of HIV-1-infected CD4$^+$ T cells in the peripheral blood and lymph nodes over time.

We found that expansions of genetically identical HIV-1 DNA sequences increased over the years of therapy at a statistically significant rate in the peripheral blood and that peripheral-blood-derived T_{EM} cells were the main contributors. This is in agreement with our previous longitudinal study, which showed that identical HIV-1 DNA sequences increased in PB-derived T_{EM} cells over 6 months of therapy in both p6-RT and *env* genomic regions (17). In contrast, the population of unique HIV-1 DNA sequences decayed in peripheral blood at a statistically significant rate, particularly in T_{EM} cells. A previous study found that an increase in the proportion of identical HIV-1 integration sites was associated with a decrease in the proportion of unique viral integration sites in PB-derived CD4$^+$ T cells during therapy (65). This supports the inverse relationship between identical and unique p6-RT HIV-1 DNA sequences within PB, particularly within the PB-derived T_{EM} cell subset. Also, it has been noted that the change in the frequency of PB-derived T_{EM} cells containing HIV-DNA was less than 2-fold during 6 months of effective therapy (17). In agreement with this previous study, we found that the HIV-1 infection frequency appeared to be stable during 3 to 18 years of therapy in T_{EM} cells sorted from peripheral blood. Taken together, our findings provide strong evidence that the overall stability of HIV-1-infected T_{EM} cells is regulated by cellular proliferation that restores T cell loss during therapy with the clonal expansion of particular T_{EM} cells containing identical HIV-1 sequences (possibly in response to an antigen) and the reduction/extinction of T_{EM} cells containing unique HIV-1 sequences. The one anatomic region that revealed a decrease in identical sequences and an increase in unique sequences was the gut; however, the findings in the gut did not reach statistical significance. This suggests that cells located in the gut are under strict immune regulation preventing their response to normal gut flora, which limits cellular proliferation and expansion of identical HIV-1 sequences (66–70).

Although there are limitations to cross-sectional analyses, this study of more than 5,000 HIV-1 single-genome sequences from 26 participants has provided important insights into the HIV-1 reservoir during 3 to 17.8 years of therapy. A longitudinal study using full-length HIV-1 sequencing and integration site analysis of individual participant samples over time would be ideal to reconfirm our findings, but it would be time-consuming and costly (43, 44, 64, 65, 71, 72). Moreover, the number of cells required to adequately sort specific T cell subsets for full-length HIV-1 sequencing and integration site analysis would limit the ability to conduct such a study using frozen peripheral blood cells and tissue samples, which most laboratories have for longitudinal samples.

In conclusion, our in-depth genetic characterization of HIV-1 within anatomic sites after 3 to 17.8 years of ART revealed several important findings. The number of HIV-1-infected memory T cells decays in all anatomic sites analyzed with early and late ART initiation. The absence of substantial increases in the pool of defective HIV-1 sequences during effective therapy in both participant cohorts suggests that the

defects are established during multiple rounds of viral replication before viral suppression rather than during effective ART. Moreover, lymph node-derived T_{EM} cells are a probable source of HIV-1 genomes capable of producing infectious virus. Importantly, the complex interplay of identical and unique HIV-1 DNA sequences indicates that cellular proliferation plays a significant role in the maintenance of persistent HIV-1, and these mechanisms are most pronounced in peripheral-blood-derived T_{EM} cells.

MATERIALS AND METHODS

Study approval. This project was approved by the institutional review board at the Western Sydney Health Department for the Westmead Institute for Medical Research (AU RED LNR/13/WMEAD/315) and the ethics review committees at the University of California San Francisco (UCSF) (10-01330/068192 and 10-02631/083640) and the Vaccine Gene Therapy Institute—Florida (VGTI-FL) (FWA 00004139). All participants provided written informed consent prior to inclusion in the study.

Participant selection. HIV-1-infected adults on effective ART for 3.0 to 17.8 years were recruited for the study at UCSF, San Francisco, CA, USA. The inclusion criteria for the study were at least 3 continuous years of therapy, with undetectable viral loads since month 6 of therapy and HIV-1 RNA at <40 copies/ml for at least 3 years. All the participants had viral loads of <40 HIV-1 RNA copies/ml during therapy except for participant 2450, whose viral load rebounded to 3,418 HIV-1 RNA copies/ml at the time of sampling (Table 1).

Clinical samples. PB samples from 26 participants infected with HIV-1 subtype B who were on long-term suppressive ART (duration, 3.0 to 17.8 years) were analyzed. Of those participants, LN and gut (gut-associated lymphoid, ileum, and/or rectum) tissues were available from 12 and 17 participants, respectively. Samples were collected from 12 participants who initiated therapy ≤6 months after infection (acute/early) and 14 participants who initiated therapy ≥1 year after HIV-1 infection (chronic) (Table 1). Note that all the clinical samples analyzed were collected after the stated duration of therapy for each participant (Table 1). The CD4 counts of all participants ranged from 339 to 1,165 cells/μl at the time of sample collection (Table 1). We completed and published in-depth longitudinal genetic characterizations of the HIV-1 of 8 participants (Tables 5 and 6) (17, 23). Here, we conducted an interpatient cross-sectional study of the impact of treatment duration on the HIV-1 reservoir, which included our published data (17, 23).

Isolation of cells from peripheral blood, lymph node, and gut tissues. The T cell subsets were isolated from peripheral blood, lymph node, and gut tissues as previously described (17, 23). Briefly, 230-ml peripheral blood samples were collected in tubes containing acid citrate dextrose as an anticoagulant or by leukapheresis. Peripheral blood mononuclear cells (PBMCs) were separated from plasma using Ficoll within 30 min after collection. Total CD4$^+$ T cells were isolated from PBMCs using magnetic negative selection (Stem Cell Technologies, Vancouver, Canada) according to the manufacturer's protocol. For isolation of cells from lymph nodes, 1/2 to 1 inguinal lymph node was removed under local anesthesia. Mechanical dissociation, followed by filtration and washing, was applied to the lymph node tissue. The isolation of cells from gut tissues and rectal and ileal biopsy specimens was accomplished using enzymatic digestion of Liberase TL or Liberase DL (Roche), respectively, in association with DNase I (Sigma) and mechanical disruption using GentleMacs (Miltenyi).

Cell sorting. Fluorescence-activated cell sorting (FACS) (FACSAria; BD Biosciences) was used to sort CD4$^+$ T cell subsets, as previously described (17, 23). CD4$^+$ T_N, T_{CM}, T_{TM}, and T_{EM} cells were sorted from peripheral blood sampled from all 26 participants (Tables 1 to 3 and Fig. 1) (17, 23). CD4$^+$ T_{SCM}, memory CXCR5$^-$ CR6$^+$ (X5$^+$ R6$^+$) and X5$^+$ R6$^-$ T cell subsets were sorted from blood samples from 6 participants whose blood samples were collected after at least 15 years of therapy (Tables 1 to 3 and Fig. 1 and 2). CD4$^+$ T_N, T_{CM}, T_{TM}, and T_{EM} cells were sorted from lymph nodes obtained from 12 participants (Tables 1 to 3 and Fig. 3) (17). Figures 1 to 3 show representative cell-sorting strategies from PB and LNs. CD4$^+$ T cells were sorted from gut tissues of 5 participants whose gut samples were collected after >15 years of therapy (participants 2115, 2518, 2046, 2013, and 2026) (Tables 1 to 3). CD4$^+$ T_N, T_{CTM}, and T_{EM} cell subsets were derived from gut samples obtained from another 12 participants using a gating strategy published previously (Tables 1 to 3) (17, 23). For peripheral blood, lymph node, and gut samples, the monoclonal antibodies used for cell sorting were CD3-V450, CD4-phycoerythrin (PE) Cy7, CD45RO-ECD, CD27-allophycocyanin (APC) AeF780, CCR7-Alexa 647, CD57-fluorescein isothiocyanate (FITC), CD45-Alexa 700, CD13-PE, and CD14-Qdot 605. Dead and live cells were distinguished with aqua-fluorescent reactive dye (ARRD). The monoclonal antibodies were obtained from Invitrogen and BD Biosciences/Pharmigen (17). For leukapheresis samples, the monoclonal antibodies used for cell sorting were CD3-Alexa 700 (clone UCHT1; BD no. 557943), CD4-APC (clone RPA-T4; BD no. 555349), CD14-V500 (clone M5E2; BD no. 561391), LIVE/DEAD aqua marker (Invitrogen no. L34957), CD45RA-BV650 (clone HI100; BioLegend no. 304135), CD45RO-peridinin chlorophyll protein (PerCP) eFluor710 (clone UCHL1; eBioscience no. 40-0457-42), CD27-APC eFluor780 (clone O323; eBioscience no. 47-0279-42), CCR7-PE CF564 (clone 15053; BD no. 562381), CD127-PE (cloneHIL-7R-M21; BD no. 557938), CD95$^-$ PECy7 (clone DX2; BD no. 561633), CCR6$^-$ PE (clone 11A9; BD no. 559562), and CXCR5-Alexa 488 (clone RF8B2; BD no. 558112). The detailed gating strategy is presented in Fig. 1 to 3. From rectal and ileal tissues, CD4$^+$ T cells were sorted from single-cell suspensions using the following combinations of antibodies: CD3-Alexa 700 (clone UCHT1; BD no. 557943), CD4-APC (clone RPA-T4; BD no. 555349), CD8-PB (clone RPA-T8; BD no. 557943), CD14-V500 (clone M5E2; BD no. 561391), and LIVE/DEAD aqua marker (Invitrogen no. L34957). The markers used for sorting each CD4$^+$ T cell subset from each anatomic site are presented in Table 3. For each cell type derived from peripheral blood, lymph nodes, and gut, 62×10^6 to 30×10^6 cells were

sorted into FACS tubes. The cells were divided and spun down in 1.5-ml Eppendorf tubes. The supernatant was removed, and the cell pellets were stored at $-80°C$ for further analysis. Postsort analysis of cellular purity was done for each cell type from peripheral blood, with means of 93.4% for T_N, 90.3% for T_{CM}, 90.9% for T_{TM}, 95.3% for T_{EM}, 97.3% for CXCR5$^-$ CCR6$^+$, and 94.9% for CXCR5$^+$ CCR6$^-$ cell subsets. The cellular purity for T_N and CD4$^+$ memory T cell subsets sorted from lymph node tissue and gut was similar to the previous report (17). The purity of CD4$^+$ T cells sorted from gut tissues was 94.9%. The cell purity for T_{SCM} cells was not assessed due to the low number of cells sorted. For HIV-1 sequencing, we analyzed an average of 5.0×10^3 to 2.4×10^6 cells in each CD4$^+$ T cell subset (Table 2).

DNA extraction. DNA was extracted from CD4$^+$ T cell subsets sorted from peripheral blood and gut samples collected from 6 participants who were on ART for at least 15 years (Table 1). Four hundred microliters of RNAzol RT (MRC, Inc.) was added to a 1.5-ml Eppendorf tube containing the cell pellet; a 0.4$\times$ RNAzol RT volume of sterile nuclease-free water (Invitrogen) was added and mixed by inversion for 15 s, followed by incubation for 15 min. The mixture was centrifuged at $16,000 \times g$ for 15 min at room temperature. The top phase was removed, and the bottom phase was used for DNA extraction. Nine hundred microliters of DNAzol (MRC, Inc.), followed by 10 μl of glycogen (20 $\mu g/\mu l$; Qiagen), was added to the bottom phase. DNA was precipitated by adding 500 μl of 200 proof ethanol (Sigma-Aldrich). The mixture was incubated for 10 min at room temperature and centrifuged at full speed for 30 min. The supernatant was removed, and the DNA pellet was washed with 75% ethanol twice. The pellet was air dried until no ethanol was visible. The pellet was dissolved in 300 μl of 8 mM NaOH (Sigma-Aldrich), followed by neutralization by adding 24 μl of 0.1 M HEPES (Gibco).

Single-proviral sequencing. The SPS assay was developed to obtain many individual intracellular HIV-1 DNA sequences from cells of participants on long-term therapy to assess the viral DNA population characteristics, diversity, evolution, and HIV infection frequencies of the cells (17, 23, 33, 73). We validated the technique using tissue culture cells with known numbers of genetically distinct HIV-1 proviruses. We applied this proven technique to quantify and genetically characterize HIV-1 DNA sequences from T cell subsets. Using advanced fluorescence-activated cell sorting, CD4$^+$ T_N, T_{SCM}, T_{CM}, T_{TM}, T_{EM}, and memory X5$^+$ R6$^-$ and X5$^-$ R6$^+$ cells were sorted from peripheral blood and lymph node tissue samples based on their cellular phenotypes (Fig. 1 and 3 and Table 3) (17, 23, 74). From gut tissues, we sorted CD4$^+$ T_N, T_{CTM}, T_{EM}, and total CD4$^+$ T cells (Table 3). Briefly, the lysate of the sorted cells or the extracted DNA was serially diluted (1:1 to 1:729). Single HIV-1 DNA molecules were amplified using primers flanking the *gag-pol* region (p6-RT). PCR amplification and sequencing of the DNA molecules in each well allowed quantification and analysis of the genetic relationship of HIV-1 DNA sequences in each infected cell subset.

Single-genome sequencing. Plasma was collected before and during ART, and HIV-1 RNA sequences were obtained using single-genome sequencing (SGS). The HIV-1 RNA sequences were compared to the HIV-1 DNA sequences identified using SPS, as described previously (75–77). At least 22 ml of plasma was used to pellet the virus. The pre-ART plasma samples were collected from 20 out of 26 participants from 4.7 years before to just prior to the initiation of treatment (Table 1) (17, 23). Eighteen out of 26 participants had plasma samples collected after ART initiation (Table 1). The on-therapy plasma samples from 5 previously reported participants were collected at 5.8 to 13.4 years of ART (17, 23). The on-therapy plasma samples from the remaining 12 participants were collected at $\leq$3 months after the initiation of therapy (Table 1). Only the sequences from the CHI group were used for the comparison of HIV-1 RNA sequences derived from pre- and on-ART plasma samples and HIV-1 DNA sequences obtained from CD4$^+$ T cell subsets sorted from different anatomic sites during therapy.

Phylogenetic analyses. Using previously described methods (17, 23), HIV-1 sequences derived from plasma and CD4$^+$ T cell subsets were phylogenetically analyzed. The SPS and SGS methods preclude the resampling of HIV-1 DNA and RNA sequences, respectively. The sequence data were generated over 3 to 4 years in two different physical locations with protocols designed to prevent cross-contamination between clinical samples and laboratory-related HIV strains. p6-RT sequences were aligned using MAFFT (78). Hypermutants (HIV-1 sequences containing G-to-A hypermutations) were identified using the Los Alamos Hypermut tool (79). Sequences containing premature stop codons, insertions/deletions causing a frameshift, and/or internal deletions were identified by manual screening and the Los Alamos quality control tool. The HIV-1 RNA sequences derived from pre-ART and early on-ART plasma samples and HIV-1 DNA sequences derived from a broad spectrum of CD4$^+$ T cell subsets sorted from PB, LN, and gut tissues were used to construct maximum-likelihood phylogenetic trees for each participant using MEGA 6 (1,000 bootstrap replicates; general time-reversible substitution model with gamma distribution and proportion of invariant sites; gamma category 4) (80). The phylogenetic trees containing all viral sequences were used to visually locate EIS (80). The genetically identical HIV-1 DNA sequences included at least 2 viral DNA sequences with 100% pairwise identity and were found in monotypic groups of a phylogenetic tree without any internal branches. Also, these identical HIV-1 DNA sequences were derived from a specific CD4$^+$ memory T cell subset sorted from an anatomic site. Maximum-likelihood phylogenetic trees at 3 different time points on ART (3.5, 10.8, and 17.3 years, respectively) were constructed to visualize the number of EIS from PB-derived intracellular HIV-1 DNA sequences at these specific time points. The unique HIV-1 DNA sequences are located on single branches within a phylogenetic tree and are genetically distinct and not genetically identical to any other viral RNA and/or DNA sequence (Fig. 14). The denominator for identical and unique HIV-1 DNA sequences, whether they are defective or not, is a single infected T cell. This is due to the fact that more than 90% of the HIV-1-infected T cells contain a single HIV-1 DNA molecule when analyzed by single-cell sequencing (33, 73). Previously published HIV-1 DNA and RNA sequence are available in GenBank (accession no. KP065816 to KP067089, KP113063 to KP113482, KP152533 to KP152580, and KP152658 to KP153066) (17).

Statistics. We used previously described maximum-likelihood methods (23) to estimate the proportion of infected cells in each cell population in each tissue type from each participant. The maximum-likelihood method was used to calculate the proportion of infected cells that would be most consistent with the observed numbers of PCR-positive wells at different dilutions with various total numbers of cells (81). We then created simpler "stand-in" data to convey equivalent information for further analysis. For each cell population from a given tissue type in a given participant, we defined stand-in data as r infected cells out of n examined, with r and n chosen so that a Poisson regression model with only an intercept term would reproduce the same estimated infection rate and confidence interval around it as obtained in the maximum-likelihood analyses. When no wells were positive, we set r equal to 0 and n equal to the total number of cells tested. We performed statistical analyses of these data using negative binomial regression, because this generalizes Poisson regression to allow additional variation. We used the SAS NL mixed procedure (SAS Institute, Cary, NC; version 9.4) to fit negative binomial regression models for each cell population from each tissue with a linear effect of years on ART, which estimates the fold effect per year on therapy. The fold effect per year on therapy is a multiplicative effect that is equivalent to the fold change in the proportion of HIV-1-infected T cells from earlier to later time points. We summarized the results across cell populations within each tissue type with geometric means. The intercept and slope coefficients from the models were used to fit curves to the number of HIV-1-infected T cells per million over years on ART, as shown in equation 1:

$$y = 1,000,000 \times e^{(k+xs)} \tag{1}$$

where y is the number of HIV-1-infected T cells per million, k is the intercept, s is the slope per year on ART, and x is the number of years on ART. This statistical calculation was applied when at least four participants contributed to the fold effect per year on ART within each T cell population and tissue.

To obtain a P value for the differences in rates of decline, we fitted a mixed-effects negative binomial regression model to the PB and LN data together, with a random intercept to account for within-person correlation of HIV-1 infection rates in PB and LN samples of the same population from the same person. These unexpectedly showed nonsignificant P values and very different estimates of the fold effect per year of ART in LNs in the AHI group, suggesting that the 4 participants with LN samples available might not be representative of the entire set of 12. The mixed-effects models used the within-person PB-LN correlations to estimate fold effects for LNs that are less susceptible to bias caused by having the data from only four participants of the AHI group (the extrapolated estimates in Fig. 4). We calculated the Pearson correlations of the logarithmically transformed proportions of T cells that were HIV-1 infected, specifically using the following formula: log[max(r,0.5)/n]. When estimated proportions were zero, we changed r to 0.5 in order to permit logarithmic calculation. We excluded gut measurements based on fewer than 1,000 cells when performing the Pearson correlations, because they were likely to be imprecise.

We defined dichotomous dependent variables for each genetic sequence of HIV-1 RNA derived from pre- and on-ART plasma samples and HIV-1 DNA sequences derived from CD4$^+$ T cell subsets sorted from peripheral blood, lymph node, and gut tissues obtained after 3 to 17.8 years of ART. These variables indicated whether a viral sequence was genetically defective, unique, or coming from EIS. An additional variable indicated whether an HIV-1 DNA sequence had an identity of 100% compared to a viral sequence derived from pre- or on-ART plasma samples (82). We applied mixed-effects logistic models with random-person effects to compare the odds of a viral sequence being genetically defective between the plasma samples, the anatomic sites, and the CD4$^+$ T cell subsets. We applied similar models to compare the odds of an HIV-1 DNA sequence being genetically identical to an HIV-1 RNA sequence derived from pre- or on-ART plasma samples, using CD4$^+$ T cell subsets (T_N, T_{CM}, T_{TM}, and T_{EM}) of peripheral blood and lymph nodes as the categorical predictor variable. For analyses involving the duration of ART, we applied mixed-effects logistic regression to estimate the odds ratio (OR) per year on ART. We plotted the proportions of defective, unique, or identical HIV-1 DNA sequences across ART duration and fitted nonlinear curves derived from mixed-effects logistic regression models using equation 2 below;

$$y = 100 \times \frac{1}{1 + e^{-(k+xs)}} \tag{2}$$

where y is the fitted percentage of defective, unique, or identical HIV-1 DNA sequences; k is the intercept; s is the log odds ratio per year on ART; and x is the number of years on ART. The fitted curves were derived from the models, with a viral sequence as the unit of analysis. Moreover, the curves account for the plotted confidence intervals, which include the number of sequences available for each time point.

Accession number(s). The HIV-1 DNA and RNA sequences have been submitted to GenBank under accession numbers MH830518 to MH834389.

ACKNOWLEDGMENTS

REFERENCES

1. Chun TW, Carruth L, Finzi D, Shen X, DiGiuseppe JA, Taylor H, Hermankova M, Chadwick K, Margolick J, Quinn TC, Kuo YH, Brookmeyer R, Zeiger MA, Barditch-Crovo P, Siliciano RF. 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature 387:183–188. https://doi.org/10.1038/387183a0.

2. Chun TW, Finzi D, Margolick J, Chadwick K, Schwartz D, Siliciano RF. 1995. In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. Nat Med 1:1284–1290. https://doi.org/10.1038/nm1295-1284.

3. Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, Lloyd AL, Nowak MA, Fauci AS. 1997. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl Acad Sci U S A 94:13193–13197. https://doi.org/10.1073/pnas.94.24.13193.

4. Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, Quinn TC, Chadwick K, Margolick J, Brookmeyer R, Gallant J, Markowitz M, Ho DD, Richman DD, Siliciano RF. 1997. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. Science 278:1295–1300. https://doi.org/10.1126/science.278.5341.1295.

5. Wong JK, Hezareh M, Gunthard HF, Havlir DV, Ignacio CC, Spina CA, Richman DD. 1997. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. Science 278:1291–1295. https://doi.org/10.1126/science.278.5341.1291.

6. Hermankova M, Siliciano JD, Zhou Y, Monie D, Chadwick K, Margolick JB, Quinn TC, Siliciano RF. 2003. Analysis of human immunodeficiency virus type 1 gene expression in latently infected resting CD4+ T lymphocytes in vivo. J Virol 77:7383–7392. https://doi.org/10.1128/jvi.77.13.7383-7392.2003.

7. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, Kovacs C, Gange SJ, Siliciano RF. 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. Nat Med 9:727–728. https://doi.org/10.1038/nm880.

8. Siliciano JD, Siliciano RF. 2005. Enhanced culture assay for detection and quantitation of latently infected, resting CD4+ T-cells carrying replication-competent virus in HIV-1-infected individuals. Methods Mol Biol 304:3–15. https://doi.org/10.1385/1-59259-907-9:003.

9. Chun TW, Justement JS, Murray D, Hallahan CW, Maenza J, Collier AC, Sheth PM, Kaul R, Ostrowski M, Moir S, Kovacs C, Fauci AS. 2010. Rebound of plasma viremia following cessation of antiretroviral therapy despite profoundly low levels of HIV reservoir: implications for eradication. AIDS 24:2803–2808. https://doi.org/10.1097/QAD.0b013e328340a239.

10. Chun TW, Engel D, Mizell SB, Ehler LA, Fauci AS. 1998. Induction of HIV-1 replication in latently infected CD4+ T cells using a combination of cytokines. J Exp Med 188:83–91. https://doi.org/10.1084/jem.188.1.83.

11. Prins JM, Jurriaans S, van Praag RM, Blaak H, van Rij R, Schellekens PT, ten Berge IJ, Yong SL, Fox CH, Roos MT, de Wolf F, Goudsmit J, Schuitemaker H, Lange JM. 1999. Immuno-activation with anti-CD3 and recombinant human IL-2 in HIV-1-infected patients on potent antiretroviral therapy. AIDS 13:2405–2410. https://doi.org/10.1097/00002030-199912030-00012.

12. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, Boucher G, Boulassel M-R, Ghattas G, Brenchley JM, Schacker TW, Hill BJ, Douek DC, Routy J-P, Haddad EK, Sékaly R-P. 2009. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. Nat Med 15:893–900. https://doi.org/10.1038/nm.1972.

13. Siliciano JD, Lai J, Callender M, Pitt E, Zhang H, Margolick JB, Gallant JE, Cofrancesco J, Jr, Moore RD, Gange SJ, Siliciano RF. 2007. Stability of the latent reservoir for HIV-1 in patients receiving valproic acid. J Infect Dis 195:833–836. https://doi.org/10.1086/511823.

14. Siliciano RF, Greene WC. 2011. HIV latency. Cold Spring Harb Perspect Med 1:a007096. https://doi.org/10.1101/cshperspect.a007096.

15. Murray AJ, Kwon KJ, Farber DL, Siliciano RF. 2016. The latent reservoir for HIV-1: how immunologic memory and clonal expansion contribute to HIV-1 persistence. J Immunol 197:407–417. https://doi.org/10.4049/jimmunol.1600343.

16. Shan L, Deng K, Gao H, Xing S, Capoferri AA, Durand CM, Rabi SA, Laird GM, Kim M, Hosmane NN, Yang HC, Zhang H, Margolick JB, Li L, Cai W, Ke R, Flavell RA, Siliciano JD, Siliciano RF. 2017. Transcriptional repro-

gramming during effector-to-memory transition renders CD4(+) T cells permissive for latent HIV-1 infection. Immunity 47:766–775 e3. https://doi.org/10.1016/j.immuni.2017.09.014.

17. von Stockenstrom S, Odevall L, Lee E, Sinclair E, Bacchetti P, Killian M, Epling L, Shao W, Hoh R, Ho T, Faria NR, Lemey P, Albert J, Hunt P, Loeb L, Pilcher C, Poole L, Hatano H, Somsouk M, Douek D, Boritz E, Deeks SG, Hecht FM, Palmer S. 2015. Longitudinal genetic characterization reveals that cell proliferation maintains a persistent HIV type 1 DNA pool during effective HIV therapy. J Infect Dis 212:596–607. https://doi.org/10.1093/infdis/jiv092.

18. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim EY, Archer J, Kosakovsky Pond SL, Chung YS, Penugonda S, Chipman JG, Fletcher CV, Schacker TW, Malim MH, Rambaut A, Haase AT, McLean AR, Wolinsky SM. 2016. Persistent HIV-1 replication maintains the tissue reservoir during therapy. Nature 530:51–56. https://doi.org/10.1038/nature16933.

19. Martinez-Picado J, Deeks SG. 2016. Persistent HIV-1 replication during antiretroviral therapy. Curr Opin HIV AIDS 11:417–423. https://doi.org/10.1097/COH.0000000000000287.

20. Buzon MJ, Massanella M, Llibre JM, Esteve A, Dahl V, Puertas MC, Gatell JM, Domingo P, Paredes R, Sharkey M, Palmer S, Stevenson M, Clotet B, Blanco J, Martinez-Picado J. 2010. HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. Nat Med 16:460–465. https://doi.org/10.1038/nm.2111.

21. Middleton T, Lim HB, Montgomery D, Rockway T, Tang H, Cheng X, Lu L, Mo H, Kohlbrenner WE, Molla A, Kati WM. 2004. Inhibition of human immunodeficiency virus type I integrase by naphthamidines and 2-aminobenzimidazoles. Antiviral Res 64:35–45. https://doi.org/10.1016/j.antiviral.2004.04.007.

22. Svarovskaia ES, Barr R, Zhang X, Pais GC, Marchand C, Pommier Y, Burke TR, Jr, Pathak VK. 2004. Azido-containing diketo acid derivatives inhibit human immunodeficiency virus type 1 integrase in vivo and influence the frequency of deletions at two-long-terminal-repeat-circle junctions. J Virol 78:3210–3222. https://doi.org/10.1128/jvi.78.7.3210-3222.2004.

23. Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, Epling L, Tan A, Ho T, Lemey P, Shao W, Hunt PW, Somsouk M, Wylie W, Douek DC, Loeb L, Custer J, Hoh R, Poole L, Deeks SG, Hecht F, Palmer S. 2013. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. Proc Natl Acad Sci U S A 110:E4987-96. https://doi.org/10.1073/pnas.1308313110.

24. Anderson JA, Archin NM, Ince W, Parker D, Wiegand A, Coffin JM, Kuruc J, Eron J, Swanstrom R, Margolis DM. 2011. Clonal sequences recovered from plasma from patients with residual HIV-1 viremia and on intensified antiretroviral therapy are identical to replicating viral RNAs recovered from circulating resting CD4+ T cells. J Virol 85:5220–5223. https://doi.org/10.1128/JVI.00284-11.

25. Wagner TA, McKernan JL, Tobin NH, Tapia KA, Mullins JI, Frenkel LM. 2013. An increasing proportion of monotypic HIV-1 DNA sequences during antiretroviral treatment suggests proliferation of HIV-infected cells. J Virol 87:1770–1778. https://doi.org/10.1128/JVI.01985-12.

26. Imamichi H, Natarajan V, Adelsberger JW, Rehm CA, Lempicki RA, Das B, Hazen A, Imamichi T, Lane HC. 2014. Lifespan of effector memory CD4+ T cells determined by replication-incompetent integrated HIV-1 provirus. AIDS 28:1091–1099. https://doi.org/10.1097/QAD.0000000000000223.

27. Kearney MF, Wiegand A, Shao W, Coffin JM, Mellors JW, Lederman M, Gandhi RT, Keele BF, Li JZ. 2016. Origin of rebound plasma HIV includes cells with identical proviruses that are transcriptionally active before stopping of antiretroviral therapy. J Virol 90:1369–1376. https://doi.org/10.1128/JVI.02139-15.

28. Simonetti FR, Sobolewski MD, Fyne E, Shao W, Spindler J, Hattori J, Anderson EM, Watters SA, Hill S, Wu X, Wells D, Su L, Luke BT, Halvas EK, Besson G, Penrose KJ, Yang Z, Kwan RW, Van Waes C, Uldrick T, Citrin DE, Kovacs J, Polis MA, Rehm CA, Gorelick R, Piatak M, Keele BF, Kearney MF, Coffin JM, Hughes SH, Mellors JW, Maldarelli F. 2016. Clonally expanded CD4+ T cells can produce infectious HIV-1 in vivo. Proc Natl Acad Sci U S A 113:1883–1888. https://doi.org/10.1073/pnas.1522675113.

29. Kulpa DA, Chomont N. 2015. HIV persistence in the setting of antiretroviral therapy: when, where and how does HIV hide? J Virus Erad 1:59–66.

30. Chun TW, Nickle DC, Justement JS, Meyers JH, Roby G, Hallahan CW, Kottilil S, Moir S, Mican JM, Mullins JI, Ward DJ, Kovacs JA, Mannon PJ, Fauci AS. 2008. Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy. J Infect Dis 197:714–720. https://doi.org/10.1086/527324.

31. Yukl SA, Gianella S, Sinclair E, Epling L, Li Q, Duan L, Choi AL, Girling V, Ho T, Li P, Fujimoto K, Lampiris H, Hare CB, Pandori M, Haase AT, Gunthard HF, Fischer M, Shergill AK, McQuaid K, Havlir DV, Wong JK. 2010. Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. J Infect Dis 202:1553–1561. https://doi.org/10.1086/656722.

32. Sathaliyawala T, Kubota M, Yudanin N, Turner D, Camp P, Thome JJ, Bickham KL, Lerner H, Goldstein M, Sykes M, Kato T, Farber DL. 2013. Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets. Immunity 38:187–197. https://doi.org/10.1016/j.immuni.2012.09.020.

33. Josefsson L, Palmer S, Faria NR, Lemey P, Casazza J, Ambrozak D, Kearney M, Shao W, Kottilil S, Sneller M, Mellors J, Coffin JM, Maldarelli F. 2013. Single cell analysis of lymph node tissue from HIV-1 infected patients reveals that the majority of CD4+ T-cells contain one HIV-1 DNA molecule. PLoS Pathog 9:e1003432. https://doi.org/10.1371/journal.ppat.1003432.

34. Gosselin A, Wiche Salinas TR, Planas D, Wacleche VS, Zhang Y, Fromentin R, Chomont N, Cohen EA, Shacklett B, Mehraj V, Ghali MP, Routy JP, Ancuta P. 2017. HIV persists in CCR6+CD4+ T cells from colon and blood during antiretroviral therapy. AIDS 31:35–48. https://doi.org/10.1097/QAD.0000000000001309.

35. Shimizu N, Tanaka A, Oue A, Mori T, Ohtsuki T, Apichartpiyakul C, Uchiumi H, Nojima Y, Hoshino H. 2009. Broad usage spectrum of G protein-coupled receptors as coreceptors by primary isolates of HIV. AIDS 23:761–769. https://doi.org/10.1097/QAD.0b013e328326cc0d.

36. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science 271:1582–1586. https://doi.org/10.1126/science.271.5255.1582.

37. Preston BD, Poiesz BJ, Loeb LA. 1988. Fidelity of HIV-1 reverse transcriptase. Science 242:1168–1171. https://doi.org/10.1126/science.2460924.

38. Rambaut A, Posada D, Crandall KA, Holmes EC. 2004. The causes and consequences of HIV evolution. Nat Rev Genet 5:52–61. https://doi.org/10.1038/nrg1246.

39. Bebenek K, Abbotts J, Roberts JD, Wilson SH, Kunkel TA. 1989. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J Biol Chem 264:16948–16956.

40. Sanchez G, Xu X, Chermann JC, Hirsch I. 1997. Accumulation of defective viral genomes in peripheral blood mononuclear cells of human immunodeficiency virus type 1-infected individuals. J Virol 71:2233–2240.

41. Okada A, Iwatani Y. 2016. APOBEC3G-mediated G-to-A hypermutation of the HIV-1 genome: the missing link in antiviral molecular mechanisms. Front Microbiol 7:2027. https://doi.org/10.3389/fmicb.2016.02027.

42. Yu Q, Konig R, Pillai S, Chiles K, Kearney M, Palmer S, Richman D, Coffin JM, Landau NR. 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nat Struct Mol Biol 11:435–442. https://doi.org/10.1038/nsmb758.

43. Hiener B, Horsburgh BA, Eden J-S, Barton K, Schlub TE, Lee E, von Stockenstrom S, Odevall L, Milush JM, Liegler T, Sinclair E, Hoh R, Boritz EA, Douek D, Fromentin R, Chomont N, Deeks SG, Hecht FM, Palmer S. 2017. Identification of genetically intact HIV-1 proviruses in specific CD4+ T cells from effectively treated participants. Cell Rep 21:813–822. https://doi.org/10.1016/j.celrep.2017.09.081.

44. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell 155:540–551. https://doi.org/10.1016/j.cell.2013.09.020.

45. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, Lai J, Strain MC, Lada SM, Hoh R, Ho YC, Richman DD, Deeks SG, Siliciano JD, Siliciano RF. 2016. Defective proviruses rapidly accumulate during acute HIV-1 infection. Nat Med 22:1043–1049. https://doi.org/10.1038/nm.4156.

46. Rusert P, Fischer M, Joos B, Leemann C, Kuster H, Flepp M, Bonhoeffer S, Gunthard HF, Trkola A. 2004. Quantification of infectious HIV-1 plasma viral load using a boosted in vitro infection protocol. Virology 326:113–129. https://doi.org/10.1016/j.virol.2004.05.022.

47. Sahu GK, Sarria JC, Cloyd MW. 2010. Recovery of replication-competent residual HIV-1 from plasma of a patient receiving prolonged, suppressive highly active antiretroviral therapy. J Virol 84:8348–8352. https://doi.org/10.1128/JVI.00362-10.

48. Vella S, Galluzzo MC, Giannini G, Pirillo MF, Andreotti M, Tomino C, Fragola V, Bucciardini R, Ricciardulli D, Binelli A, Weimer LE, Floridia M. 2000. Plasma HIV-1 copy number and in vitro infectivity of plasma prior to and during combination antiretroviral treatment. Antiviral Res 47:189–198. https://doi.org/10.1016/s0166-3542(00)00107-8.

49. Zhu W, Jiao Y, Lei R, Hua W, Wang R, Ji Y, Liu Z, Wei F, Zhang T, Shi X, Wu H, Zhang L. 2011. Rapid turnover of 2-LTR HIV-1 DNA during early stage of highly active antiretroviral therapy. PLoS One 6:e21081. https://doi.org/10.1371/journal.pone.0021081.

50. Gandhi RT, McMahon DK, Bosch RJ, Lalama CM, Cyktor JC, Macatangay BJ, Rinaldo CR, Riddler SA, Hogg E, Godfrey C, Collier AC, Eron JJ, Mellors JW, ACTG A5321 Team. 2017. Levels of HIV-1 persistence on antiretroviral therapy are not associated with markers of inflammation or activation. PLoS Pathog 13:e1006285. https://doi.org/10.1371/journal.ppat.1006285.

51. Laskey SB, Pohlmeyer CW, Bruner KM, Siliciano RF. 2016. Evaluating clonal expansion of HIV-infected cells: optimization of PCR strategies to predict clonality. PLoS Pathog 12:e1005689. https://doi.org/10.1371/journal.ppat.1005689.

52. Imamichi H, Dewar RL, Adelsberger JW, Rehm CA, O'Doherty U, Paxinos EE, Fauci AS, Lane HC. 2016. Defective HIV-1 proviruses produce novel protein-coding RNA species in HIV-infected patients on combination antiretroviral therapy. Proc Natl Acad Sci U S A 113:8783–8788. https://doi.org/10.1073/pnas.1609057113.

53. Pollack RA, Jones RB, Pertea M, Bruner KM, Martin AR, Thomas AS, Capoferri AA, Beg SA, Huang S-H, Karandish S, Hao H, Halper-Stromberg E, Yong PC, Kovacs C, Benko E, Siliciano RF, Ho Y-C. 2017. Defective HIV-1 proviruses are expressed and can be recognized by cytotoxic T lymphocytes, which shape the proviral landscape. Cell Host Microbe 21:494–506.e4. https://doi.org/10.1016/j.chom.2017.03.008.

54. Shete A, Thakar M, Singh DP, Gangakhedkar R, Gaikwad A, Pawar J, Paranjape R. 2012. Short communication: HIV antigen-specific reactivation of HIV infection from cellular reservoirs: implications in the settings of therapeutic vaccinations. AIDS Res Hum Retroviruses 28:835–843. https://doi.org/10.1089/aid.2010.0363.

55. Buzon MJ, Martin-Gayo E, Pereyra F, Ouyang Z, Sun H, Li JZ, Piovoso M, Shaw A, Dalmau J, Zangger N, Martinez-Picado J, Zurakowski R, Yu XG, Telenti A, Walker BD, Rosenberg ES, Lichterfeld M. 2014. Long-term antiretroviral treatment initiated at primary HIV-1 infection affects the size, composition, and decay kinetics of the reservoir of HIV-1-infected CD4 T cells. J Virol 88:10056–10065. https://doi.org/10.1128/JVI.01046-14.

56. Soriano-Sarabia N, Bateson RE, Dahl NP, Crooks AM, Kuruc JD, Margolis DM, Archin NM. 2014. Quantitation of replication-competent HIV-1 in populations of resting CD4+ T cells. J Virol 88:14070–14077. https://doi.org/10.1128/JVI.01900-14.

57. Ramratnam B, Mittler JE, Zhang L, Boden D, Hurley A, Fang F, Macken CA, Perelson AS, Markowitz M, Ho DD. 2000. The decay of the latent reservoir of replication-competent HIV-1 is inversely correlated with the extent of residual viral replication during prolonged anti-retroviral therapy. Nat Med 6:82–85. https://doi.org/10.1038/71577.

58. Locci M, Havenar-Daughton C, Landais E, Wu J, Kroenke MA, Arlehamn CL, Su LF, Cubas R, Davis MM, Sette A, Haddad EK, International AVIPCPI, Poignard P, Crotty S. 2013. Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses. Immunity 39:758–769. https://doi.org/10.1016/j.immuni.2013.08.031.

59. Breitfeld D, Ohl L, Kremmer E, Ellwart J, Sallusto F, Lipp M, Forster R. 2000. Follicular B helper T cells express CXC chemokine receptor 5, localize to B cell follicles, and support immunoglobulin production. J Exp Med 192:1545–1552. https://doi.org/10.1084/jem.192.11.1545.

60. Zeng M, Southern PJ, Reilly CS, Beilman GJ, Chipman JG, Schacker TW, Haase AT. 2012. Lymphoid tissue damage in HIV-1 infection depletes naive T cells and limits T cell reconstitution after antiretroviral therapy. PLoS Pathog 8:e1002437. https://doi.org/10.1371/journal.ppat.1002437.

61. Zanini F, Puller V, Brodin J, Albert J, Neher RA. 2017. In vivo mutation rates and the landscape of fitness costs of HIV-1. Virus Evol 3:vex003. https://doi.org/10.1093/ve/vex003.

62. Cuevas JM, Geller R, Garijo R, López-Aldeguer J, Sanjuán R. 2015. Extremely high mutation rate of HIV-1 in vivo. PLoS Biol 13:e1002251. https://doi.org/10.1371/journal.pbio.1002251.

63. Huang SH, Ren Y, Thomas AS, Chan D, Mueller S, Ward AR, Patel S, Bollard CM, Cruz CR, Karandish S, Truong R, Macedo AB, Bosque A, Kovacs C, Benko E, Piechocka-Trocha A, Wong H, Jeng E, Nixon DF, Ho YC, Siliciano RF, Walker BD, Jones RB. 2018. Latent HIV reservoirs exhibit inherent resistance to elimination by CD8+ T cells. J Clin Invest 128: 876–889. https://doi.org/10.1172/JCI97555.

64. Lee GQ, Orlova-Fink N, Einkauf K, Chowdhury FZ, Sun X, Harrington S, Kuo H-H, Hua S, Chen H-R, Ouyang Z, Reddy K, Dong K, Ndung'u T, Walker BD, Rosenberg ES, Yu XG, Lichterfeld M. 2017. Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. J Clin Invest 127:2689–2696. https://doi.org/10.1172/JCI93289.

65. Cohn LB, Silva IT, Oliveira TY, Rosales RA, Parrish EH, Learn GH, Hahn BH, Czartoski JL, McElrath MJ, Lehmann C, Klein F, Caskey M, Walker BD, Siliciano JD, Siliciano RF, Jankovic M, Nussenzweig MC. 2015. HIV-1 integration landscape during latent and active infection. Cell 160: 420–432. https://doi.org/10.1016/j.cell.2015.01.020.

66. Ribeiro A, Heimesaat MM, Bereswill S. 2017. Changes of the intestinal microbiome-host homeostasis in HIV-infected individuals—a focus on the bacterial gut microbiome. Eur J Microbiol Immunol 7:158–167. https://doi.org/10.1556/1886.2017.00016.

67. Pabst O, Mowat AM. 2012. Oral tolerance to food protein. Mucosal Immunol 5:232–239. https://doi.org/10.1038/mi.2012.4.

68. Angin M, Klarenbeek PL, King M, Sharma SM, Moodley ES, Rezai A, Piechocka-Trocha A, Toth I, Chan AT, Goulder PJ, Ndung'u T, Kwon DS, Addo MM. 2014. Regulatory T cells expanded from HIV-1-infected individuals maintain phenotype, TCR repertoire and suppressive capacity. PLoS One 9:e86920. https://doi.org/10.1371/journal.pone.0086920.

69. Klatt NR, Funderburg NT, Brenchley JM. 2013. Microbial translocation, immune activation, and HIV disease. Trends Microbiol 21:6–13. https://doi.org/10.1016/j.tim.2012.09.001.

70. Angin M, Kwon DS, Streeck H, Wen F, King M, Rezai A, Law K, Hongo TC, Pyo A, Piechocka-Trocha A, Toth I, Pereyra F, Ghebremichael M, Rodig SJ, Milner DA, Jr, Richter JM, Altfeld M, Kaufmann DE, Walker BD, Addo MM. 2012. Preserved function of regulatory T cells in chronic HIV-1 infection despite decreased numbers in blood and tissue. J Infect Dis 205: 1495–1500. https://doi.org/10.1093/infdis/jis236.

71. Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, Spindler J, Ferris AL, Mellors JW, Kearney MF, Coffin JM, Hughes SH. 2014. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. Science 345:179–183. https://doi.org/10.1126/science.1254194.

72. Wagner TA, McLaughlin S, Garg K, Cheung CY, Larsen BB, Styrchak S, Huang HC, Edlefsen PT, Mullins JI, Frenkel LM. 2014. HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. Science 345:570–573. https://doi.org/10.1126/science.1256304.

73. Josefsson L, King MS, Makitalo B, Brannstrom J, Shao W, Maldarelli F, Kearney MF, Hu WS, Chen J, Gaines H, Mellors JW, Albert J, Coffin JM, Palmer SE. 2011. Majority of CD4+ T cells from peripheral blood of HIV-1-infected individuals contain only one HIV DNA molecule. Proc Natl Acad Sci U S A 108:11199–11204. https://doi.org/10.1073/pnas.1107729108.

74. Josefsson L, Eriksson S, Sinclair E, Ho T, Killian M, Epling L, Shao W, Lewis B, Bacchetti P, Loeb L, Custer J, Poole L, Hecht FM, Palmer S. 2012. Hematopoietic precursor cells isolated from patients on long-term suppressive HIV therapy did not contain HIV-1 DNA. J Infect Dis 206:28–34. https://doi.org/10.1093/infdis/jis301.

75. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, Rock D, Falloon J, Davey RT, Jr, Dewar RL, Metcalf JA, Hammer S, Mellors JW, Coffin JM. 2005. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. J Clin Microbiol 43:406–413. https://doi.org/10.1128/JCM.43.1.406-413.2005.

76. Kearney M, Maldarelli F, Shao W, Margolick JB, Daar ES, Mellors JW, Rao V, Coffin JM, Palmer S. 2009. Human immunodeficiency virus type 1 population genetics and adaptation in newly infected individuals. J Virol 83:2715–2727. https://doi.org/10.1128/JVI.01960-08.

77. Kearney M, Palmer S, Maldarelli F, Shao W, Polis MA, Mican J, Rock-Kress D, Margolick JB, Coffin JM, Mellors JW. 2008. Frequent polymorphism at drug resistance sites in HIV-1 protease and reverse transcriptase. AIDS 22:497–501. https://doi.org/10.1097/QAD.0b013e3282f29478.

78. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/mst010.

79. Rose PP, Korber BT. 2000. Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. Bioinformatics 16:400–401. https://doi.org/10.1093/bioinformatics/16.4.400.

80. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.

81. Rosenbloom DI, Elliott O, Hill AL, Henrich TJ, Siliciano JM, Siliciano RF. 2015. Designing and interpreting limiting dilution assays: general principles and applications to the latent reservoir for human immunodeficiency virus-1. Open Forum Infect Dis 2:ofv123. https://doi.org/10.1093/ofid/ofv123.

82. De Scheerder M-A, Vrancken B, Dellicour S, Schlub T, Lee E, Shao W, Rutsaert S, Verhofstede C, Kerre T, Malfait T, Hemelsoet D, Coppens M, Dhondt A, De Looze D, Vermassen F, Lemey P, Palmer S, Vandekerckhove L. 2019. HIV rebound is predominantly fueled by genetically identical viral expansions from diverse reservoirs. Cell Host Microbe 26: 347–358.e7. https://doi.org/10.1016/j.chom.2019.08.003.