Title

CXCR4-Using HIV Strains Predominate in Naive and Central Memory CD4+ T Cells in People Living with HIV on Antiretroviral Therapy: Implications for How Latency Is Established and Maintained

Journal

Journal of Virology, 94(6)

ISSN

0022-538X

Authors

Roche, Michael
Tumpach, Carolin
Symons, Jori
et al.

Publication Date

2020-02-28

DOI

10.1128/jvi.01736-19

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

CXCR4-Using HIV Strains Predominate in Naive and Central Memory CD4$^+$ T Cells in People Living with HIV on Antiretroviral Therapy: Implications for How Latency Is Established and Maintained

Michael Roche,[a,b] Carolin Tumpach,[a,b] Jori Symons,[a] Matthew Gartner,[b] Jenny L. Anderson,[a] Gabriela Khoury,[a] Kieran Cashin,[b] Paul U. Cameron,[a,c] Melissa J. Churchill,[b] Steven G. Deeks,[d] Paul R. Gorry,[b] Sharon R. Lewin[a,c]

[a]The Peter Doherty Institute for Infection and Immunity, University of Melbourne and Royal Melbourne Hospital, Melbourne, Australia
[b]School of Health and Biomedical Sciences, RMIT University, Bundoora, Australia
[c]Department of Infectious Diseases, Monash University and Alfred Hospital, Melbourne, Australia
[d]Department of Medicine, University of California, San Francisco, California, USA

ABSTRACT HIV can persist in people living with HIV (PLWH) on antiretroviral therapy (ART) in multiple CD4$^+$ T cell subsets, including naive cells, central memory (CM) cells, transitional (TM) cells, and effector memory (EM) cells. Since these cells express different levels of the viral coreceptors CXCR4 and CCR5 on their surface, we sought to determine whether the HIV envelope protein (Env) was genotypically and phenotypically different between CD4$^+$ T cell subsets isolated from PLWH on suppressive ART ($n = 8$). Single genome amplification for the HIV *env* gene was performed on genomic DNA extracts from different CD4$^+$ T cell subsets. We detected CXCR4-using (X4) strains in five of the eight participants studied, and in these participants, the prevalence of X4 strains was higher in naive CD4$^+$ T cells than in the memory subsets. Conversely, R5 strains were mostly found in the TM and EM populations. Identical sets of *env* sequences, consistent with clonal expansion of some infected cells, were more frequent in EM cells. These expanded identical sequences could also be detected in multiple CD4$^+$ T cell subsets, suggesting that infected cells can undergo T cell differentiation. These identical sequences largely encoded intact and functional Env proteins. Our results are consistent with a model in which X4 HIV strains infect and potentially establish latency in naive and CM CD4$^+$ T cells through direct infection, in addition to maintenance of the reservoir through differentiation and proliferation of infected cells.

IMPORTANCE In people living with HIV (PLWH) on suppressive ART, latent HIV can be found in a diverse range of CD4$^+$ T cells, including quiescent naive and central memory cells that are typically difficult to infect *in vitro*. It is currently unclear how latency is established in these cells *in vivo*. We show that in CD4$^+$ T cells from PLWH on suppressive ART, the use of the coreceptor CXCR4 was prevalent among viruses amplified from naive and central memory CD4$^+$ T cells. Furthermore, we found that expanded numbers of identical viral sequences were most common in the effector memory population, and these identical sequences were also found in multiple different CD4$^+$ T cell subsets. Our results help to shed light on how a range of CD4$^+$ T cell subsets come to harbor HIV DNA, which is one of the major barriers to eradicating the virus from PLWH.

KEYWORDS CCR5, CXCR4, HIV, latency, tropism

The principal barrier to a cure for human immunodeficiency virus (HIV) infection is the persistence of long-lived and proliferating latently infected CD4$^+$ T cells in people living with HIV (PLWH) on antiretroviral therapy (ART) (1). Although HIV persis-

Citation Roche M, Tumpach C, Symons J, Gartner M, Anderson JL, Khoury G, Cashin K, Cameron PU, Churchill MJ, Deeks SG, Gorry PR, Lewin SR. 2020. CXCR4-using HIV strains predominate in naive and central memory CD4$^+$ T cells in people living with HIV on antiretroviral therapy: implications for how latency is established and maintained. J Virol 94:e01736-19. https://doi.org/10.1128/JVI.01736-19.

Editor Frank Kirchhoff, Ulm University Medical Center

Address correspondence to Sharon R. Lewin, sharon.lewin@unimelb.edu.au.

Received 9 October 2019
Accepted 9 December 2019
Accepted manuscript posted online 18 December 2019
Published 28 February 2020

TABLE 1 Participant characteristics

Participant	Age (yrs)	No. of yrs on ART[a]	CD4 (cells/μl)	Nadir CD4 (cells/μl)
1	58	11.2	489	87
2	65	6.0	735	227
3	61	10.1	672	547
4	50	12.7	696	147
5	60	13.3	1,145	314
6	70	10.0	866	665
7	59	4.2	416	75
8	43	2.7	408	134

[a]ART, antiretroviral therapy.

tence on ART was initially thought to primarily involve resting, central memory (CM) CD4$^+$ T cells, HIV can also persist in naive, stem memory (TSCM), transitional memory (TM), effector memory (EM), and terminally differentiated (TD) CD4$^+$ T cells (2, 3). Latently infected cells can persist on ART due to the long half-life of some cells, such as naive and memory CD4$^+$ T cells, and also through proliferation leading to clones of infected cells (4–6). However, the role of cellular maturation and differentiation in maintaining the latent reservoir remains unclear. Defining how the latent reservoir is established and how it persists during ART is critical to the design of approaches to cure HIV infection.

The cellular tropism of HIV is determined by the viral envelope glycoproteins (Env), and viruses are broadly classified based on which coreceptor Env engages; R5 viruses use CCR5 exclusively, X4 viruses use CXCR4 exclusively, and R5X4 viruses can use either coreceptor. The CD4$^+$ T cell subsets that can be persistently infected during ART express different cell surface levels of CCR5 and CXCR4, with CCR5 being highest on EM, then TM, CM, and, finally, naive CD4$^+$ T cells, where CCR5 is barely detectable. The expression of CXCR4 is the opposite that of CCR5, with the highest levels on naive CD4$^+$ T cells, then CM, TM, and EM cells (7–11). If latency is established primarily through direct infection, one might therefore expect a relative enrichment of X4 variants in naive and CM cells, with R5 variants enriched in TM and EM cells, in people on long-term ART.

In this study, we sought to characterize Envs isolated from CD4$^+$ T cell subsets from blood and tissue of PLWH on ART to define the relationship between viruses in each subset based on coreceptor usage and Env function. In a cohort of PLWH who initiated ART with a low nadir CD4$^+$ T cell count, we found that X4 strains were frequently detected, largely in naive and CM T cells, while R5 strains were detected in EM and TM cells. Expanded numbers of identical *env* sequences were most common in the EM cell population, but these identical *env* sequences were also found in multiple different CD4$^+$ T cell subsets. Our results are consistent with a model in which X4 HIV strains infect and potentially establish latency in naive and CM CD4$^+$ T cells through a direct mechanism, while R5 HIV strains preferentially infect more differentiated (TM and EM) cells. Furthermore, consistent with previous reports, clonal expansion of more differentiated infected cells (TM and EM) and potentially cellular differentiation contribute to HIV persistence on ART.

RESULTS

***env* sequences cluster according to coreceptor tropism, not cell of origin.** To investigate the role of the HIV Env in HIV persistence on ART, we used single genome amplification and sequencing to characterize *env* genes from blood and tissue collected from PLWH on ART who were all male (Table 1). The demographics of this cohort have been previously described (12–14). Briefly, the inclusion criteria for the study were being on ART with a plasma viral load of $<$50 HIV RNA copies per ml for at least two and a half years. Blood, lymph node (LN) excisional biopsy specimens, and rectal biopsy specimens were collected while participants were on ART, and for some participants, plasma was available prior to ART. The frequency of latently infected cells is highest in lymph node and the gastrointestinal tract in PLWH on suppressive ART (14, 15), and

TABLE 2 *env* amplimers generated

Particant	No. of amplimers on indicated cell type or tissue										
	Naive	Central memory $X3^+$ $R6^+$	Central memory $X3^+$ $R6^-$	Central memory $X3^-$ $R6^+$	Central memory $X3^-$ $R6^-$	Transitional memory	Effector memory $PD1^+$	Effector memory $PD1^-$	Lymph node	Rectal tissue	Plasma (pretherapy)
1	—[a]	12	9	7	6	15	4	7	NA[b]	1	NA
2	19	14	21	9	15	11	17	10	NA	16	22
3	10	7	4	8	2	9	5	5	NA	6	13
4	13	9	12	13	9	19	NA	12	4	14	NA
5	19	10	4	8	2	9	9	16	12	4	NA
6	9	6	12	17	20	16	13	14	1	9	NA
7	6	7	—[a]	5	9	13	1	6	NA	NA	NA
8	5	14	11	11	6	14	NA	15	NA	NA	NA

[a]—, *env* sequences could not be recovered from this sample.
[b]NA, not available.

biopsy specimens were included to determine if tissue and blood contained different *env* sequences.

We were able to amplify and sequence at least 5 *env* sequences for the majority of subsets from each sample (61/71 [85%]) (Table 2). To understand the contribution of defective *env* sequences, we first identified defective *env* sequences. The proportion of hypermutated sequences detected was <10% in 6/8 participants, while participants 1 and 2 had relatively higher levels of hypermutated sequences (24.5% and 14.5%, respectively) (Fig. 1A). The proportion of defective *env* sequences (determined by the

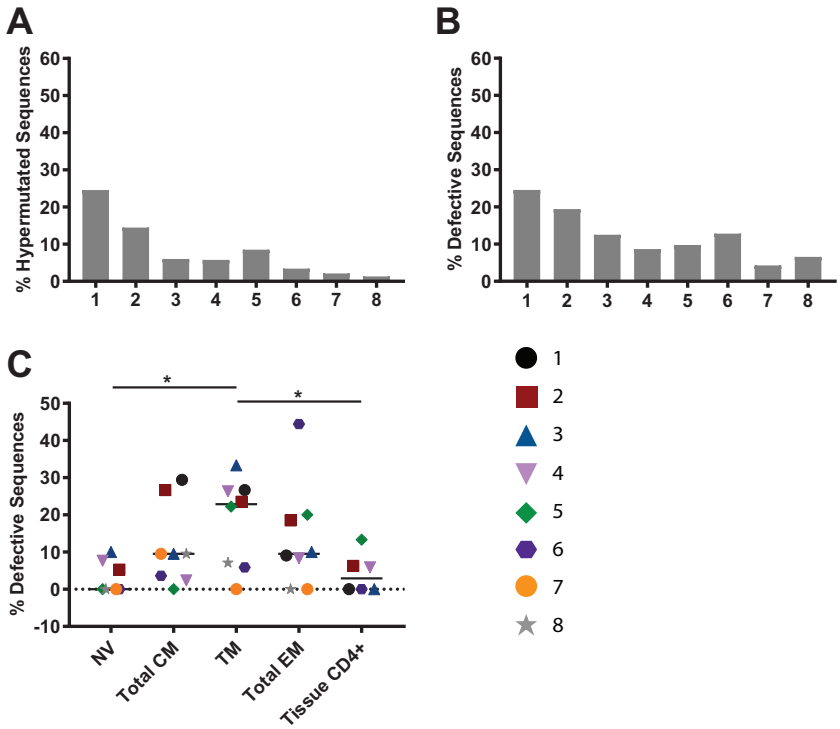

FIG 1 Distribution of hypermutated and defective *env* sequences. (A and B) The proportions of hypermutated sequences (determined by hypermut 2.0 program) (A) and defective sequences (hypermutated sequences, sequences containing stop codons or deletions of receceptor binding sites) (B) as a percentage of total sequences are shown for all participants (identified as participants 1 to 8). (C) Frequency of defective sequences in $CD4^+$ T cell subsets from blood and $CD4^+$ T cells isolated from either rectal or lymph node tissue (tissue $CD4^+$). Scatterplots were constructed with each participant represented by a different color and symbol, and horizontal lines indicate the medians. Comparisons were made using the Wilcoxon signed-rank test. *, $P < 0.05$. NV, naive; CM, central memory; TM, transitional memory; EM, effector memory.

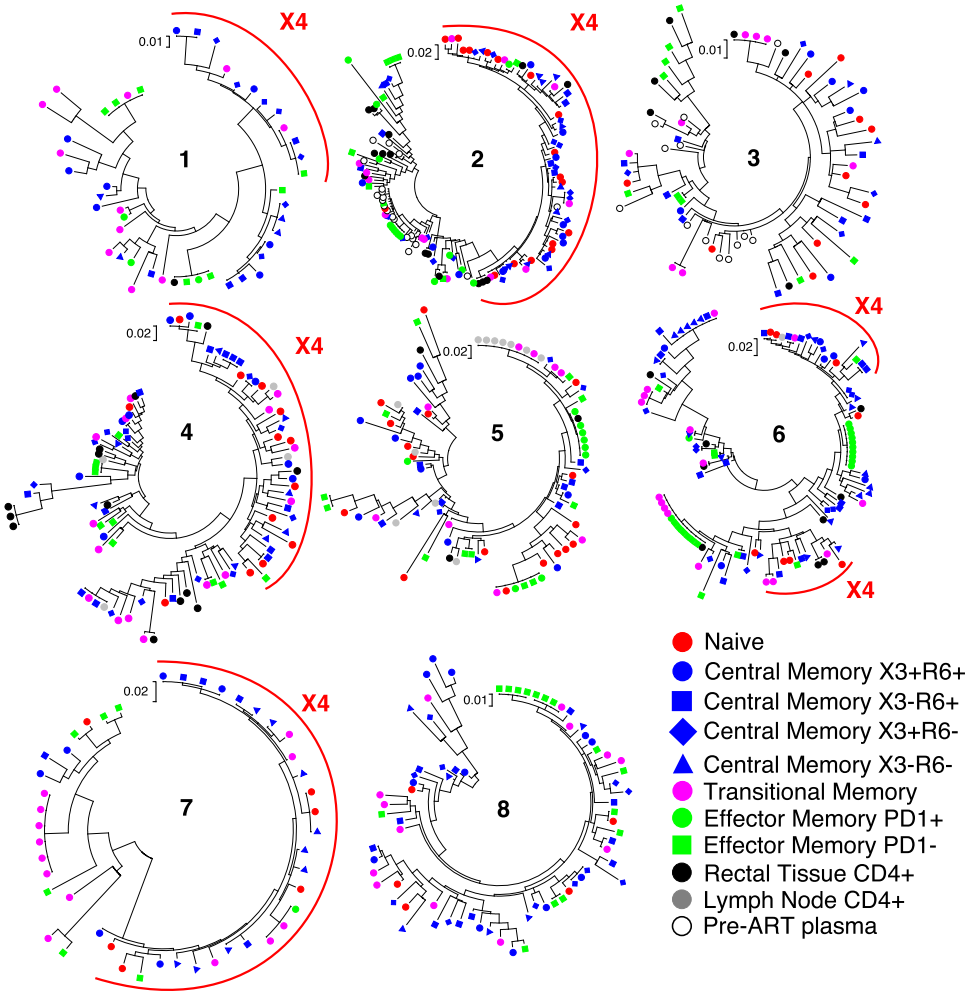

FIG 2 Phylogenetic analysis of *env* sequences. Shown are maximum likelihood trees for *env* sequences isolated from peripheral blood (naive, central memory, transitional memory, and effector memory), rectal tissue (total CD4$^+$), and lymph node (total CD4$^+$) collected from PLWH on ART. For two participants, *env* sequences were analyzed from plasma collected prior to the initiation of ART (3 months for participant 2 and 29 months for participant 3). Red semicircles indicate CXCR4-using *env* clusters. The scale is number of nucleotide substitutions per site.

presence of deleterious hypermutations, stop codons, frameshifts, or deletions encompassing receptor binding sites) was also higher in participants 1 and 2 (Fig. 1B), which was related to the higher level of hypermutated sequences isolated from these participants. The proportions of defective sequences were similar across all CD4$^+$ T cell subsets in blood and total CD4$^+$ T cells in tissue, with the exception of an increased proportion of defective sequences in TM cells from blood compared to that in naive cells from blood and tissue CD4$^+$ T cells (Fig. 1C).

To determine if *env* sequences were genotypically distinct based on the cell of origin, we constructed phylogenetic trees of *env* sequences for each individual after the removal of hypermutated variants. Phylogenetic analysis showed that *env* sequences did not cluster based on the cell of origin (Fig. 2). Consistent with this finding, we did not observe compartmentalization between *env* sequences isolated from different CD4$^+$ T cell subsets in the majority of participants, nor was compartmentalization detected between blood- and tissue-derived *env* sequences (Table 3).

Although we did not observe genetic similarity in *env* sequences based on cell of origin, we hypothesized that compartmentalization may occur based on *env* function or phenotype. We therefore determined the coreceptor tropism of each *env* sequence using the geno2pheno coreceptor prediction algorithm (16). CXCR4-predicted variants

TABLE 3 Panmixia tests[a]

Comparison	Result for indicated participant							
	1	2	3	4	5	6	7	8
Pre-ART vs. naive	—	**<0.0001**	0.9900	—	—	—	—	—
Pre-ART vs. CM	—	**<0.0001**	0.9500	—	—	—	—	—
Pre-ART vs. TM	—	0.0002	0.0290	—	—	—	—	—
Pre-ART vs. EM	—	**<0.0001**	0.0120	—	—	—	—	—
Pre-ART vs. LN	—	—	—	—	—	—	—	—
Pre-ART vs. RT	—	**<0.0001**	0.0047	—	—	—	—	—
Naive vs. CM	—	0.4500	0.8700	0.2100	0.0750	0.0400	0.3100	0.8200
Naive vs. TM	—	0.0066	0.0400	0.0110	0.0780	0.0071	0.0440	0.8200
Naive vs. EM	—	**<0.0001**	0.0022	0.0020	0.0130	0.0042	0.0850	0.0530
Naive vs. LN	—	—	—	0.2700	0.0004	—	—	—
Naive vs. RT	—	**<0.0001**	0.0620	0.0008	0.5200	0.0130	—	—
CM vs. TM	0.2500	0.0970	0.0089	0.5500	0.0300	0.0019	0.0006	0.4900
CM vs. EM	0.0690	**<0.0001**	0.0011	0.0025	0.0005	**<0.0001**[b]	0.0015	**<0.0001**[b]
CM vs. LN	—	—	—	0.8500	0.0002	—	—	—
CM vs. RT	—	0.0001	0.0027	0.0031	0.5000	0.0350	—	—
TM vs. EM	0.7800	0.0370	0.0160	0.0630	0.1100	0.0100	0.0660	0.0230
TM vs. LN	—	—	—	1.0000	0.1500	—	—	—
TM vs. RT	—	0.1600	0.2500	0.0770	0.3100	0.1800	—	—
EM vs. LN	—	—	—	0.1500	**<0.0001**[b]	—	—	—
EM vs. RT	—	0.0043	0.2200	0.0190	0.4100	0.0170	—	—
LN vs. RT	—	—	—	0.3800	0.1000	—	—	—

[a]Panmixia tests were performed as previously described (45). A *P* value of <0.0001 was considered significant. Statistically significant results are underlined and in boldface. —, panmixia test was not performed due to insufficient *env* sequences, as described in Materials and Methods.

[b]Significance was lost when identical sequences were removed from analysis (participant 5 EM versus LN, $P = 0.83$; participant 6 CM versus EM, $P = 0.0038$; participant 8 CM versus EM, $P = 0.57$).

were detected in 5 of 8 participants, with a mean (range) percentage frequency of 46.2% (28.9 to 62.2%) of total *env* sequences. These five individuals harboring X4 variants had a lower median (range) $CD4^+$ T cell count nadir (147 [75 to 665] cells/μl) than those with R5-only variants (547 [134 to 314] cells/μl). The levels of integrated HIV DNA did not differ between the two groups (Fig. 3).

Visualization of X4 variants on phylogenetic trees demonstrated two patterns; the X4 *env* sequences either formed monophyletic clusters distinct from R5 *env* sequences in the same participants (participants 1, 2, 4, and 7) (Fig. 2), or they formed multiple

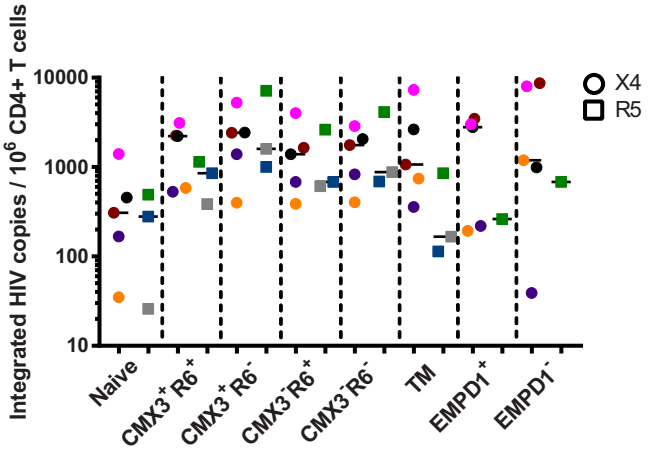

FIG 3 Integrated HIV DNA in $CD4^+$ T cell subsets. Integrated HIV DNA copies per 10^6 $CD4^+$ T cell subsets (naive, CM, TM, and EM) were determined previously (12, 13). CM cells were further differentiated by expression of CXCR3 and CCR6, and EM cells were differentiated by expression of PD1. Scatterplots were constructed with each participant represented by a different color, and the horizontal lines indicate the medians. Individuals harboring X4 *env* sequences are represented by circles, and those harboring R5-only *env* sequences are represented by squares. Black, participant 1; red, participant 2; blue, participant 3; pink, participant 4; green, participant 5; purple, participant 6; orange, participant 7; gray, participant 8. Note that integrated HIV DNA for EM^{PD1-} was not available for participants 3 and 8.

distinct clusters within a larger R5 population (participant 6) (Fig. 2). The two distinct X4 clusters in participant 6 differed primarily in the V1/V2 region rather than the V3 loop (data not shown) and may indicate two separate events when X4 viruses emerged. Our data demonstrate *env* clustering in PLWH on suppressive ART based not on cell of origin but upon predicted coreceptor phenotype.

CXCR4-using Envs are prevalent in naive and central memory CD4$^+$ T cells. Analysis of the phylogenetic trees from the five participants with X4 HIV showed that *env* sequences isolated from naive T cells were mostly excluded from branches of R5 viruses, while *env* sequences isolated from EM T cells were mostly excluded from branches of X4 viruses. We therefore next investigated the proportion of X4 *env* sequences within each CD4$^+$ T cell subset from these five participants. A gradient of CXCR4 usage was observed across the spectrum of cell differentiation. The proportion of total *env* sequences that were X4 tropic was highest in naive CD4$^+$ T cells (median, 86.06%), then CM (47.62%), TM (33.33%), and EM (13.04%) cells (Fig. 4A). To confirm this gradient of coreceptor tropism at an individual sequence level, we analyzed the false-positive rate (FPR) of each sequence from each CD4$^+$ T cell subset. The FPR is a qualitative value provided by the geno2pheno program which defines the probability of classifying a CCR5-predicted virus falsely as a CXCR4-predicted variant. The FPR was lower in *env* sequences derived from naive cells than from CM ($P = 0.0074$), TM ($P \leq 0.0001$), and EM ($P \leq 0.0001$) T cells. These results are consistent with a reduction in frequency of X4-tropic viruses as cells become more differentiated (Fig. 4B). We did not detect any differences in the proportion of X4 *env* sequences or FPR within subsets of CM cells that express either CXCR3 or CCR6 (Fig. 4C and D), nor were any differences observed within subsets of EM cells that express programmed cell death protein 1 (PD1) (Fig. 4E and F). The frequency of X4 *env* sequences in total CD4$^+$ T cells from rectal tissue was low (33.3%), similar to that in TM and EM cells from blood (Fig. 4G). Conversely, the frequency of X4 *env* sequences in CD4$^+$ T cells from lymph node was relatively high, although limited conclusions can be made given the relatively small sample size (Fig. 4G; Table 2). The FPR was significantly higher for *env* sequences isolated from rectal tissue ($P = 0.0427$) (Fig. 4H). To ensure that our analysis was not biased by clonal expansion or proliferation, we reanalyzed the data using only unique sequences or a set of identical sequences if the sequence was derived exclusively from one cell subpopulation (and therefore was counted as a single sequence). Similar results were obtained when using all sequences or only unique sequences (data not shown). The coreceptor tropism of *env* sequences closely matched the receptor expression of CCR5 on these cell populations, with EM cells expressing the highest level of CCR5 while naive cells expressed negligible levels of CCR5 (Fig. 5).

To confirm the predictions of coreceptor usage based on the geno2pheno algorithm, we verified the coreceptor tropism of 108 *env* sequences from six participants using a phenotypic assay. We created Env-pseudotyped luciferase reporter viruses and tested their ability to mediate viral entry into U87-CD4/CCR5 and U87-CD4/CXCR4 cells. The majority of *env* sequences could mediate entry into cells when pseudotyped onto Env-deficient luciferase reporter viruses and were therefore functional (89.8%) (Table 4). We saw high concordance between the genotypic prediction and phenotypic assessment of coreceptor usage (98.9% concordance), with the majority of *env* sequences classified by genotype as CXCR4 using capable of using both CCR5 and CXCR4 coreceptors in the phenotypic assay.

Together, our results demonstrate a clear segregation of *env* sequences from different CD4$^+$ T cell subsets based on coreceptor tropism. Importantly, the coreceptor usage was concordant with the known coreceptor expression on the cells, with the highest proportion of X4 *env* sequences isolated from naive and CM cells, which express high levels of CXCR4 but minimal CCR5 (7). X4 *env* sequences were isolated at a far lower frequency from TM and EM cells, which express comparatively higher levels of CCR5.

Naive and central memory cell derived CCR5-using Envs do not utilize CCR5 more efficiently. Given that R5 *env* sequences were also identified from naive and CM

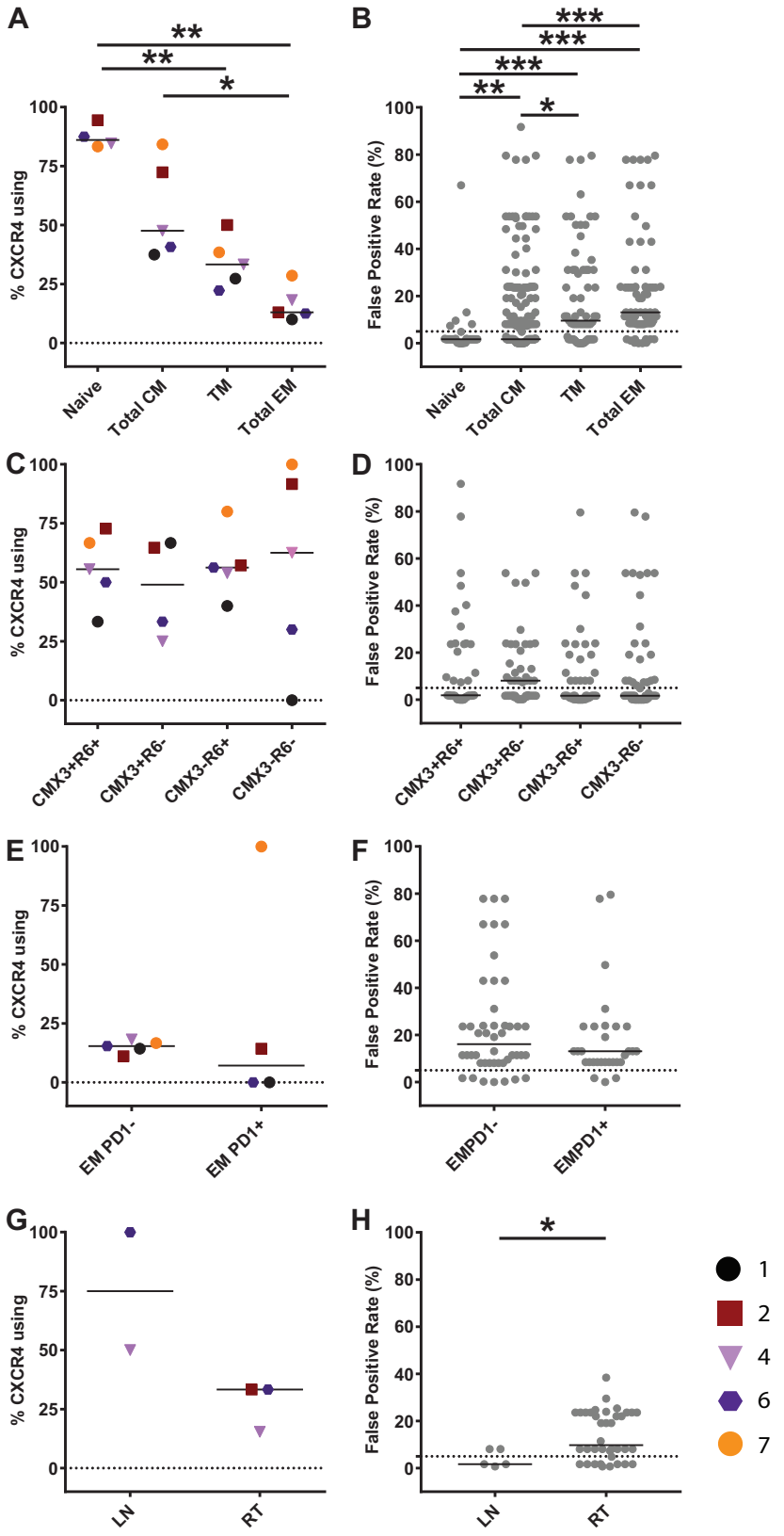

FIG 4 CXCR4-using *env* sequences are predominant in naive and central memory CD4$^+$ T cells. Shown are the frequencies of *env* sequences predicted to use X4 as a percentage of total sequences (left) and *env* false-positive rate (FPR; determined by geno2pheno prediction algorithm) (right) in naive, CM, TM, and EM CD4$^+$ T cells from peripheral blood (A and B), central memory CD4$^+$ T cells from blood based on expression of CXCR3 and CCR6 (C and D), effector memory CD4$^+$ T cells from blood based on

(Continued on next page)

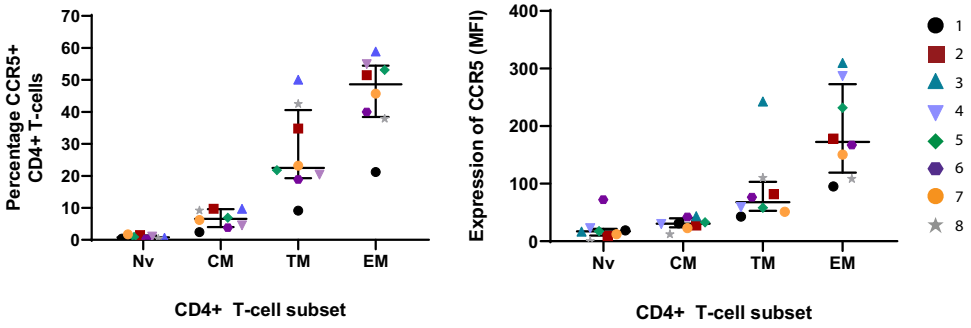

FIG 5 CCR5 expression on sorted CD4$^+$ T cell subsets. The frequencies of CCR5-positive cells (left) and the mean fluorescence intensities (MFI) of CCR5 staining (right) of naive, central memory, transitional memory and effector memory CD4$^+$ T cell subsets were determined as previously described (12, 13). Scatterplots were constructed with each participant represented by a different color and symbol, and lines indicate the medians.

cells, we hypothesized that these *env* sequences would have enhanced Env-CCR5 interactions due to the low CCR5 expression on naive and CM cells. We created Env-pseudotyped luciferase reporter viruses from R5 *env* sequences isolated from different CD4$^+$ T cell subsets and assessed their sensitivity to inhibition by the CCR5 antagonist maraviroc (MVC), as a means of comparing the efficiencies of CCR5 usage between Envs (17). Envs with enhanced CCR5 usage would be expected to be less sensitive to reductions in available CCR5 due to increases in MVC concentration and will thus have higher MVC 50% and 90% inhibitory concentrations (IC_{50} and IC_{90}) (18). We created Env-pseudotyped reporter viruses from a range of representative R5 Envs amplified from CD4$^+$ T cells from six participants (Table 4) and determined their sensitivity to inhibition by MVC. We observed no difference in MVC sensitivity as measured by IC_{50} (Fig. 6A) and IC_{90} (Fig. 6B) between R5 *env* sequences isolated from naive, CM, TM, and EM cells. These data support the conclusion that *env* sequences from naive and CM cells compared to *env* sequences from TM and EM cells do not have enhanced interactions with CCR5.

CXCR4 usage is associated with enhanced tropism for resting CD4$^+$ T cells. To further understand the importance of coreceptor usage in the direct infection of target T cells, we assessed the ability of Envs isolated from different CD4$^+$ T cell subsets to mediate infection into resting and activated CD4$^+$ T cells. We pseudotyped green fluorescent protein (GFP) reporter viruses with Envs isolated from different CD4$^+$ T cell subsets. The expression of GFP (an indicator of overall levels of infection) was higher following infection of activated CD4$^+$ T cells than for resting CD4$^+$ T cells, as expected (Fig. 7A). We did not observe any differences between infection of resting and activated CD4$^+$ T cells based on whether Envs were isolated from naive, CM, TM, or EM CD4$^+$ T cells. Interestingly, a subset of Envs were able to infect both resting and activated CD4$^+$ T cells with similar efficiencies. We hypothesized that these Envs were capable of using CXCR4, so we analyzed infection efficiency based on the Env coreceptor usage. R5 Envs were restricted to infection of activated CD4$^+$ T cells (Fig. 7B and C). Conversely, X4 Envs were able to infect both resting and activated CD4$^+$ T cells (Fig. 7B and C). To confirm that the differences in infection by X4 and R5 using Envs of resting and activated CD4$^+$ T cells were due to the expression of CXCR4 and CCR5, we infected both resting and activated CD4$^+$ T cells with six of the dual-tropic Envs we isolated in

FIG 4 Legend (Continued)
expression of PD1 and (E and F), and total CD4$^+$ T cells isolated from lymph node (LN) and rectal tissue (RT) (G and H). Scatterplots were constructed with each participant represented by a different color and symbol, and lines indicate the medians. For panels A, C, E, and G, comparisons were made using one-way analysis of variance (ANOVA) with Tukey's *post hoc* test for multiple comparisons. For panels B, D, F, and H, the dotted line indicates the FPR cutoff for CXCR4 usage (5%). Comparisons were made using a Kruskal-Wallis test with Dunn's *post hoc* test for multiple comparisons. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

TABLE 4 Genotypic and phenotypic characterization of coreceptor tropism by selected *env* sequences

Participant	Subset	Clone	Tropism (genotype)	FPR[a]	Viral entry (RLU)[b]		Tropism (phenotype)
					U87-CD4/CCR5	U87-CD4/CXCR4	
1	TM	16	X4	1.3	++	++	R5X4
	CM X3$^+$ R6$^-$	9	X4	1.1	++	++	R5X4
	CM X3$^-$ R6$^+$	2	X4	1.1	++	++	R5X4
	TM	6	R5	50.2	+++	−	R5
	TM	15	X4	1.1	−	−	NF[c]
	EM PD1$^+$	6	R5	79.5	+++	−	R5
	CM X3$^-$ R6$^-$	8	R5	23.9	−	−	NF
	EM PD1$^-$	9	R5	77.8	+++	−	R5
	EM PD1$^-$	3	R5	23.9	+++	−	R5
	TM	18	X4	1.7	+	+	R5X4
	CM X3$^+$ R6$^-$	3	R5	23.9	+++	−	R5
	TM	17	R5	50.2	+++	−	R5
	EM PD1$^-$	1	R5	23.9	+++	−	R5
	CM X3$^+$ R6$^+$	3	X4	1.1	++	++	R5X4
	EM PD1$^+$	4	R5	77.8	+++	−	R5
2	CM X3$^+$ R6$^-$	12	R5	15.4	+++	−	R5
	EM PD1$^-$	4	R5	23.6	+++	−	R5
	TM	3	R5	23.6	+++	−	R5
	EM PD1$^+$	17	R5	13.1	−	−	NF
	EM PD1$^-$	8	R5	23.6	+	−	R5
	TM	1	X4	1.7	++	+++	R5X4
	NV	20	X4	1.7	++	+++	R5X4
	NV	26	X4	1.9	++	+++	R5X4
	NV	8	X4	1.7	++	+++	R5X4
	NV	11	R5	13.1	+++	−	R5
	CM X3$^+$ R6$^+$	14	X4	1.9	−	++	X4
	CM X3$^-$ R6$^+$	9	X4	1.7	++	+++	R5X4
	EM PD1$^+$	6	R5	13.1	+++	−	R5
	TM	4	X4	1.7	+++	+++	R5X4
	TM	6	R5	29.5	−	−	NF
	EM PD1$^-$	2	R5	20.8	+++	−	R5
	NV	14	X4	1.7	−	−	NF
	CM X3$^+$ R6$^-$	21	X4	1.7	++	+++	R5X4
	CM X3$^-$ R6$^-$	10	X4	1.7	++	+++	R5X4
3	TM	8	R5	17.8	++	−	R5
	TM	10	R5	17.8	++	−	R5
	EM PD1$^-$	2	R5	21.4	++	−	R5
	EM PD1$^+$	1	R5	17.8	+++	−	R5
	CM X3$^+$ R6$^-$	6	R5	17.8	+++	−	R5
	CM X3$^-$ R6$^-$	2	R5	17.8	++	−	R5
	NV	7	R5	17.8	++	−	R5
	NV	11	R5	17.8	++	−	R5
	TM	11	R5	17.8	+++	−	R5
	EM PD1$^-$	1	R5	14.7	+++	−	R5
	CM X3$^+$ R6$^+$	5	R5	17.8	++	−	R5
	CM X3$^-$ R6$^+$	10	R5	17.8	−	−	NF
	NV	1	R5	17.8	++	−	R5
	CM X3$^+$ R6$^+$	4	R5	17.8	++	−	R5
	NV	8	R5	17.8	++	−	R5
	TM	12	R5	23.9	++	−	R5
	NV	4	R5	17.8	+	−	R5
4	TM	10	R5	8.1	+++	−	R5
	CM X3$^-$ R6$^+$	1	X4	1.7	++	++	R5X4
	CM X3$^-$ R6$^-$	5	X4	1.7	++	++	R5X4
	EM PD1$^-$	14	R5	8.1	++	−	R5
	NV	1	X4	1.7	−	−	NF
	NV	8	X4	1.7	++	++	R5X4
	CM X3$^+$ R6$^+$	10	R5	8.1	+++	−	R5
	TM	7	R5	8.1	−	−	NF
	TM	20	X4	1.7	+++	++	R5X4

(Continued on next page)

TABLE 4 (Continued)

Participant	Subset	Clone	Tropism (genotype)	FPR[a]	Viral entry (RLU)[b]		Tropism (phenotype)
					U87-CD4/CCR5	U87-CD4/CXCR4	
	EM PD1$^-$	8	R5	8.1	++	−	R5
	EM PD1$^-$	15	X4	1.7	++	+	R5X4
	CM X3$^+$ R6$^+$	6	X4	1.7	++	++	R5X4
	TM	5	R5	8.1	+++	−	R5
	CM X3$^+$ R6$^-$	7	R5	8.1	+++	−	R5
	NV	3	R5	9.6	+++	−	R5
	TM	8	R5	8.1	++	−	R5
	EM PD1$^-$	11	R5	9.6	+++	−	R5
	NV	4	X4	1.7	++	++	R5X4
	NV	2	X4	1.7	−	−	NF
	EM PD1$^-$	7	R5	43.0	+	−	R5
5	NV	1	R5	17.1	+++	−	R5
	TM	2	R5	33.9	+	−	R5
	CM X3$^+$ R6$^+$	6	R5	17.1	+++	−	R5
	NV	2	R5	33.9	+++	−	R5
	TM	7	R5	23.6	+++	−	R5
	CM X3$^+$ R6$^+$	8	R5	78.1	+++	−	R5
	CM X3$^-$ R6$^-$	2	R5	23.6	+++	−	R5
	EM PD1$^+$	1	R5	43.8	+++	−	R5
	NV	7	R5	21.3	++	−	R5
	TM	10	R5	23.6	+++	−	R5
	EM PD1$^-$	19	R5	78.8	+++	−	R5
	CM X3$^+$ R6$^-$	4	R5	43.8	+++	−	R5
	NV	6	R5	45.7	+++	−	R5
	CM X3$^-$ R6$^+$	9	R5	33.9	−	−	NF
	EM PD1$^-$	8	R5	33.9	+++	−	R5
	EM PD1$^-$	9	R5	43.3	+++	−	R5
	NV	10	R5	33.9	+++	−	R5
6	EM PD1$^+$	4	R5	11.4	+++	−	R5
	TM	7	R5	53.8	+++	−	R5
	TM	20	R5	11.4	+++	−	R5
	EM PD1$^-$	3	R5	53.8	+++	−	R5
	EM PD1$^-$	8	R5	23.6	++	−	R5
	NV	14	X4	0.7	−	+++	X4
	TM	15	R5	31.1	++	−	R5
	NV	9	R5	7.4	+++	−	R5
	NV	13	X4	4.8	+++	−	R5
	TM	4	X4	0.7	−	−	NF
	NV	10	X4	0.2	+	+++	R5X4
	TM	12	R5	8.5	+++	−	R5
	EM PD1$^-$	16	X4	0.2	−	+++	X4
	EM PD1$^+$	5	R5	31.1	+++	−	R5
	CM X3$^+$ R6$^+$	12	X4	0.2	−	++	X4
	CM X3$^+$ R6$^-$	4	X4	0.2	−	++	X4
	CM X3$^-$ R6$^+$	20	X4	0.2	+	+++	R5X4
	CM X3$^-$ R6$^-$	4	R5	53.8	+++	−	R5
	CM X3$^-$ R6$^-$	18	R5	7.4	+++	−	R5
	CM X3$^+$ R6$^+$	11	R5	11.4	+++	−	R5
Controls	NL4.3		X4	0.5	−	+++	X4
	AD8		R5	35.3	+++	−	R5
	Mock		−	−	−	−	−

[a]FPR, false-positive rate.
[b]RLU, relative light fluorescence: 0 to 10^4 (−), 10^4 to 10^5 (+), 10^5 to 10^6 (++), or 10^6 to 10^7 (+++).
[c]NF, nonfunctional.

the presence of MVC and AMD3100 (AMD), an antagonist of CXCR4 (19). Infection with dual-tropic Env-pseudotyped viruses in the presence of MVC had minimal effect on virus entry in both resting and activated CD4$^+$ T cells, suggesting that CXCR4 was the primary receptor used by dual-tropic Envs (Fig. 7D). Infection in the presence of the CXCR4 inhibitor AMD led to a 47% (±3.4%) reduction in viral entry in activated CD4$^+$

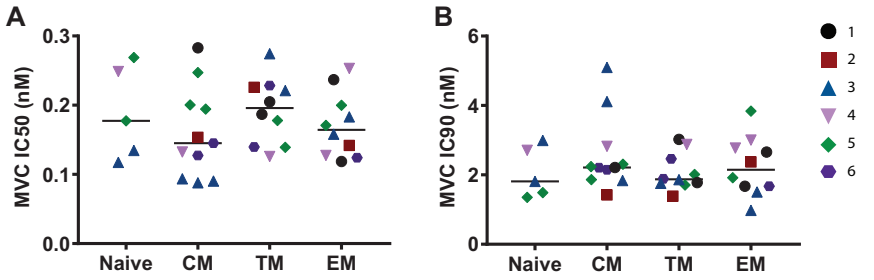

FIG 6 Sensitivity of Envs from different CD4$^+$ T cell subsets to maraviroc. The IC_{50} (A) and IC_{90} (B) for sensitivity to maraviroc using reporter viruses pseudotyped with selected Envs isolated from naive, CM, TM, and EM CD4$^+$ T cells from peripheral blood were determined. Scatterplots were constructed with each participant represented by a different color and symbol, and the horizontal lines indicate the medians. Data points represent the means from 3 independent experiments.

T cells; however, entry in resting CD4$^+$ T cells was greatly reduced. These data indicate that both CCR5 and CXCR4 are available on activated CD4$^+$ T cells for viral entry, while only CXCR4 is available on resting CD4$^+$ T cells.

Identical expanded *env* sequences are more common in effector memory CD4$^+$ T cells but are detected in all T cell subsets. In each participant we observed multiple sets of identical *env* sequences, which we assume were from different cells, due to the use of single genome amplification. We defined an identical sequence as a sequence with at least one other identical sequence within the same individual. Overall, 27.9% of all *env* sequences were identical. When identical sequences were stratified by cell subset of origin, the proportion of total sequences that were identical was higher in EM cells than in naive and CM cells from blood (Fig. 8A). We did not detect differences in the proportion of identical sequences when CM cells were stratified based on the expression of CXCR3 and CCR6 (Fig. 8B), or when EM cells were stratified based on the expression of PD1 (Fig. 8C) given work showing that HIV was enriched in cells expressing PD1 (12), or between sequences derived from lymph nodes and rectal tissue (Fig. 8D), although our results may have been influenced by low numbers of *env* amplified from some of the subsets (Table 2). Despite an increase in the proportion of identical sequences in EM T cells, we did not detect a reduction in overall *env* diversity in this subset (Fig. 8E).

We identified a total of 55 sets of identical sequences within the eight participants studied and then determined which subsets and subpopulations were harboring identical sequences. Interestingly, 58.2% of the sets of identical sequence contained sequences derived from at least two different CD4$^+$ T cell subsets or subpopulations (Fig. 9). For example, the same sets of identical *env* sequences (labeled set 4 from participant 5) (Fig. 9) were isolated from naive, TM, EM^{PD1+}, and EM^{PD1-} CD4$^+$ T cells. Some sets of identical *env* sequences were identified in only CM subpopulations (21.8%) and EM subpopulations (20.0%). In addition, we detected 11 sets of identical *env* sequences in CD4$^+$ T cells isolated from tissue (lymph node or rectal). In these sets of identical sequences from tissue, 7/11 (63.6%) were found in both tissue- and blood-derived CD4$^+$ T cells. Interestingly, in the 5 participants harboring dual-tropic variants, we found that 63.4% of identical sequence sets were R5, while 36.6% were X4. This was consistent with the overall level of R5 (53.8%) and X4 (46.2%) strains in these participants. Of the 108 *env* sequences tested for functionality, 20 were from identical sets of sequences, and 88 were unique sequences. Of the *env* sequences from identical sets of sequences, 90% (18/20) produced functional Env proteins in our pseudotyping assay (Fig. 9). Of the unique *env* sequences, 92% (81/88) were functional. These results suggest that in our cohort of PLWH on ART, expanded HIV *env* sequences can be either R5 or X4 and can harbor intact and functional envelopes at a frequency similar to that of unique sequences.

These data based on subgenomic sequencing of *env* are consistent with proliferation of infected CD4$^+$ T cells on ART, as reported in previous studies (4, 6, 20), and with

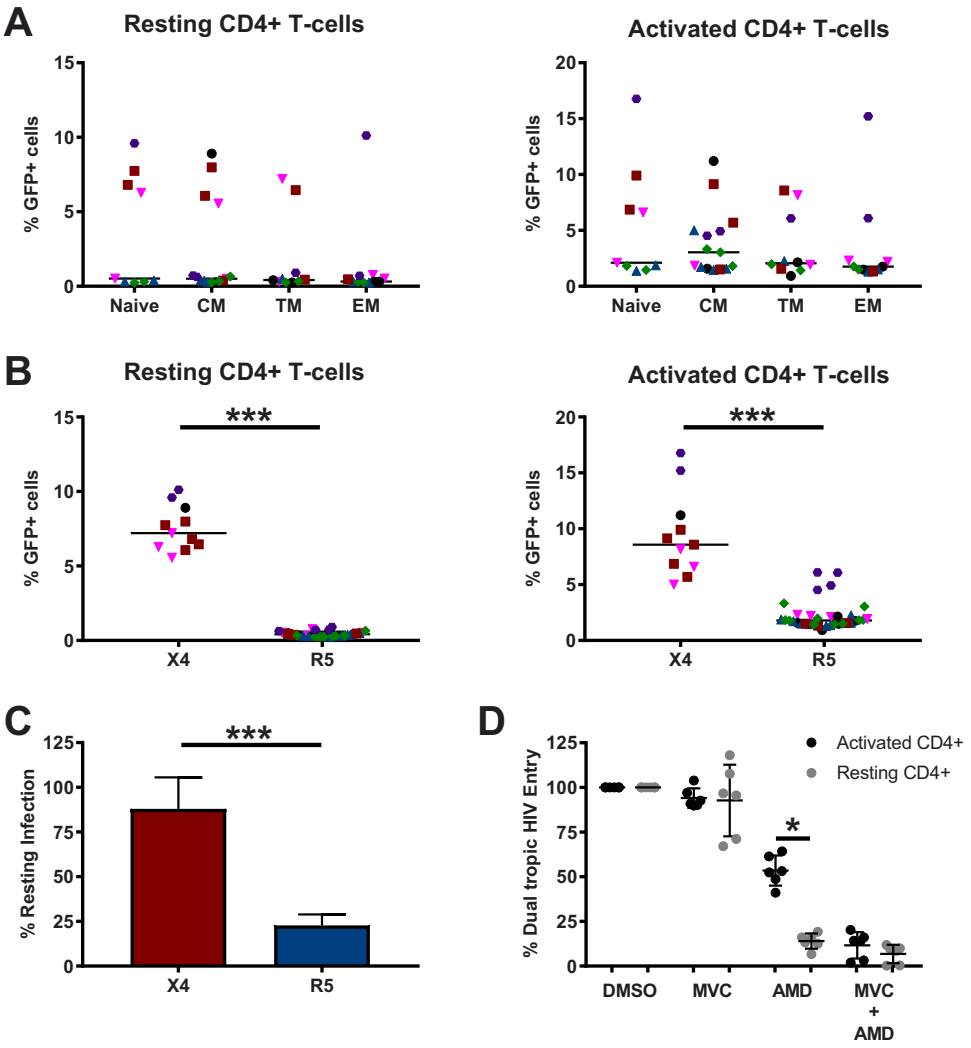

FIG 7 CXCR4-using Envs efficiently infect resting CD4$^+$ T cells. Env-pseudotyped GFP reporter viruses were used to infect resting and activated CD4$^+$ T cells from healthy donors. Envs were stratified based on cell of origin (A) and coreceptor tropism (B). The data are presented by scatterplots, with lines indicating the median. Data points represent the means from infections in 4 donors. The ratio of infection in resting CD4$^+$ T cells to activated CD4$^+$ T cells is indicated in panel C. The data are represented by bars, with error bars indicating the SEs using X4 Envs ($n = 11$) and R5 Envs ($n = 34$). (D) Inhibition of six dual-tropic Envs in activated and resting CD4$^+$ T cells by maraviroc (MVC) and AMD-3100 (AMD). The data are represented by scatterplots, with lines indicating the means and SEs from dual-tropic Envs ($n = 6$; participant 2 clones CM10, NV20, and TM4 and participant 4 clones CM5, NV4, and TM20). Black, participant 1; red, participant 2; blue, participant 3; pink, participant 4; green, participant 5; purple, participant 6. For panels A, B, and C, comparisons were made with the Mann-Whitney U test. For panel D, comparisons were made with the Wilcoxon signed-rank test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

the notion that this occurs most commonly in EM cells. However, the presence of identical expanded *env* sequences in different CD4$^+$ T cell subsets and in cells expressing different levels of CXCR4 or CCR5 is consistent with maturation or differentiation of infected cells also occurring on ART. Furthermore, the presence of identical expanded *env* sequences amplified from CD4$^+$ T cells from blood and tissue is consistent with trafficking of infected expanded clones between blood and tissue compartments.

DISCUSSION

Our results demonstrate that in PLWH on ART who initiate ART with a low nadir CD4 T cell count, HIV *env* genes are not uniform across different T cell subsets and in blood and tissue, as assessed by coreceptor tropism. CXCR4 usage was most common in *env* genes isolated from naive and CM CD4$^+$ T cells, consistent with the relatively high

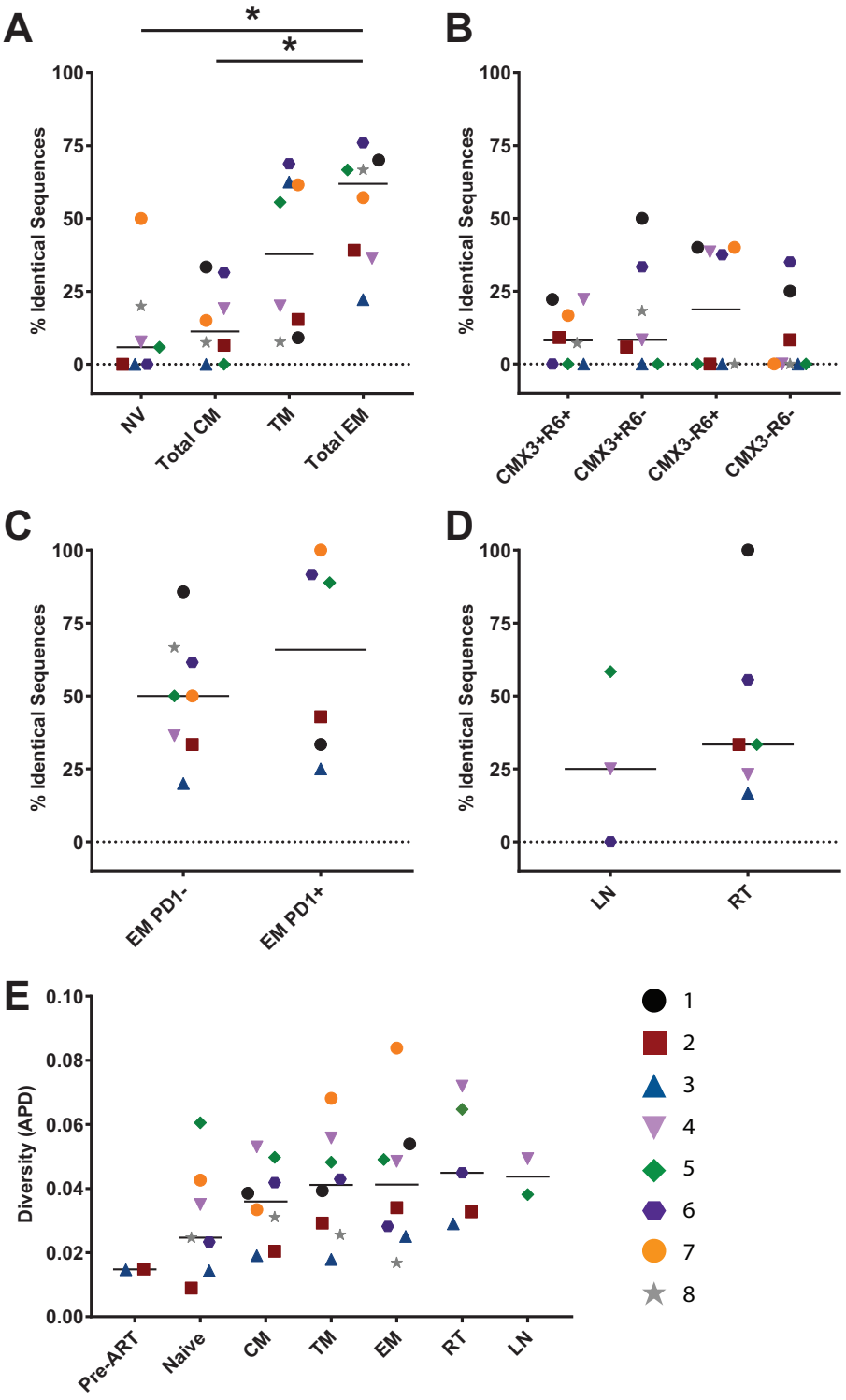

FIG 8 Identical sets of *env* sequences are most common in effector memory CD4$^+$ T cells. The percentages of identical sets of *env* sequences (two or more sequences with the same V1 to V5 sequence) of total *env* sequences were determined in naive, CM, TM, and EM CD4$^+$ T cells from peripheral blood (A), central memory CD4$^+$ T cells from blood based on expression of CXCR3 and CCR6 (B), effector memory CD4$^+$ T cells from blood based on expression of PD1 (C), and total CD4$^+$ T cells isolated from lymph node and rectal tissue (D). (E) *env* diversity in CD4$^+$ T cells subsets was determined by calculation of average pairwise distance. Scatterplots were constructed, and horizontal lines indicate the medians. Comparisons were made using the Wilcoxon signed-rank test. APD, average pairwise distance. *, $P < 0.05$.

FIG 9 Identical and functional *env* sequences are found in multiple CD4$^+$ T cell subsets. Sets of identical *env* sequences from each participant were assigned a unique number, and the distribution of the identical *env* sequences in T cell subsets and/or tissue is indicated by a circle (sequence detected) or the absence of a circle (sequence not detected). The number of times the sequence was observed is represented by the size of the circle. Envelope protein function was determined following infection of U87-CD4/CCR5 and U87-CD4/CXCR4 cells with luciferase reporter viruses pseudotyped with a limited number of Envs that were detected from identical sets of *env* sequences. Functionality is indicated by boxes on the left of the graphs, with a green check indicating a functional Env protein and a red cross indicating a nonfunctional protein.

expression of CXCR4 on these cells. Conversely, *env* from TM and EM cells was predominantly R5 tropic, consistent with the relatively high expression of CCR5 on these cells. Furthermore, infection with pseudotyped viruses containing R5 Env occurred in activated CD4$^+$ T cells only, while pseudotyped viruses containing X4 Env could infect both resting and activated CD4$^+$ T cells. This observation is likely due to the reduced availability of CCR5 on resting CD4$^+$ T cells. We observed sets of identical

env sequences in different T cell subsets; however, sets of identical sequences were concentrated in the EM cells, suggesting that clonal expansion occurred most commonly in EM cells. The majority of these expanded *env* sequences also encoded functional Env proteins.

We detected X4 Envs in five of eight participants. This is not a surprising finding considering all participants were treated during chronic infection and the majority had low nadir CD4 T cell counts (Table 1). The presence of relatively high proportions of X4 *env* in naive and CM $CD4^+$ T cells raises the question of whether CXCR4 usage is required for efficient infection of these cell subsets, even in participants who initiate ART early following infection. As X4 strains typically emerge during later stages of disease, this might suggest that early ART leads to a relative sparing of the naive and CM T cell compartments from infection. A previous study compared the frequencies of infected naive and CM $CD4^+$ T cells in PLWH on ART following treatment in acute and chronic infections but found no difference in the level of HIV DNA, although the authors did not specifically assess the frequency of X4 HIV strains (21). Studies in a cohort of PLWH who can successfully control virus to low levels after cessation of ART (post treatment controllers) suggested that early ART may limit the frequency of infected naive and CM $CD4^+$ T cells (22, 23). Others have shown that when X4 variants are present, there is a higher level of HIV DNA and inducible replication-competent provirus in naive $CD4^+$ T cells—both on and off ART (24–27). Furthermore, *in vitro* experiments have shown that infection of naive T cells primarily occurs with X4 strains (28–30). Given the clear detection of X4 *env* in cells that express CXCR4, and that only viruses pseudotyped with X4 Env led to efficient infection of resting $CD4^+$ T cells, our data support a model in which latency is established in naive and CM cells following direct infection of these cells. Conversely, R5 *env* was predominantly amplified from cells that express CCR5, such as EM and TM cells. In addition, pseudotyped viruses with R5 Env were largely only able to infect activated $CD4^+$ T cells.

In the three participants who harbored only R5 *env*, HIV DNA was detected in all T cell subsets (Fig. 3), and when R5 Envs isolated from naive and CM subsets were used to pseudotype reporter viruses, the virus did not use CCR5 more efficiently or gain the capacity to infect resting $CD4^+$ T cells. R5 Envs typically do not mediate infection of naive cells *in vitro*, and this may be related to the very low levels of CCR5 expression on these cells (7, 8, 10, 28, 29). We also detected negligible levels of CCR5 on naive $CD4^+$ T cells. Consistent with our findings, previous studies have also detected R5 *env* from naive T cells isolated from PLWH both on and off ART (3, 25).

While X4 *env* was more commonly found in CM cells in our study, R5 *env* was also detected in these cells. CM $CD4^+$ T cells express relatively low levels of CCR5, but when activated, they increase their expression of CCR5 (31). Interestingly, low CCR5 expression in some nonhuman primates infected with simian immunodeficiency virus (SIV) has been associated with a relative sparing of infection of the CM population (32). In studies of Env strains resistant to CCR5 antagonists, enhanced CCR5 usage was associated with tropism for CM (10, 11). However, in contrast to these studies, we did not find any variation in CCR5 usage of Envs isolated from CM.

We propose several explanations for how infected naive and CM T cells are established on ART. The first is through direct infection of naive and CM with X4 variants due to the high levels of CXCR4 on these cells (7). The second is that an activated $CD4^+$ T cell is infected with an X4 or R5 strain, survives, and then reverts to become either an infected naive or CM $CD4^+$ T cell (33). Reversion of EM to CM is a process called dedifferentiation, which occurs through epigenetic reprograming, at least in $CD8^+$ T cells (34). Whether this occurs in $CD4^+$ T cells remains unclear. The third is the direct infection of cells transitioning from EM to CM with R5 using viruses due to a temporary upregulation of CCR5 (31). Finally, direct infection of memory $CD4^+$ T cells with an R5 virus is possible but in the presence of an additional stimulus, such as dendritic cell contact, as we have previously shown *in vitro* (35, 36). Although these pathways could explain R5 infection of CM $CD4^+$ T cells, they would not explain our findings of R5 infection of naive $CD4^+$ T cells.

Our sequence analysis showed that identical sets of *env* sequences were more frequently detected in EM cells than in naive and CM CD4$^+$ T cells. Identical sets of *env* sequences were also identified in multiple T cell subsets. In fact, the majority of identical sets of *env* sequences (59.3%) were detected in at least two different T cell subsets. Our data are consistent with proliferation of infected cells being an important mechanism for viral persistence during ART (4, 5, 37), particularly in EM cells (20). Our observation of identical sets of *env* sequences found across multiple different T cell subsets is in agreement with other studies with participants on ART (20) and with HIV controllers (38). These data suggest that infected cells can undergo maturation and differentiation and survive. Our observation that identical sequences can be found in subpopulations of CM cells based on CXCR3 and CCR6 expression further supports this model, as the findings are consistent with infection of a precursor population prior to differentiation into Th1 (CXCR3$^+$) or Th17 (CCR6$^+$) cells. Interestingly, the bulk of identical sets of *env* sequences we detected contained intact *env* sequences that were functional. Indeed, recent evidence has shown that infected cells with intact proviruses can proliferate without producing virus (4).

Our study had some limitations. First, we recruited participants with low nadir CD4$^+$ T cell counts. As a consequence, the detection of CXCR4-using variants in five of the eight participants was not surprising but may not translate to PLWH who initiate ART with high CD4 counts. Second, as this was a cross-sectional study, we were unable to make any conclusions in relation to the relative contributions of clonal expansion, differentiation, and direct infection to how HIV persists on ART. Follow-up studies using longitudinal sampling will be important to better understand these mechanisms. Third, in the absence of full-length proviral sequencing (39), we were unable to determine whether the *env* genes described here are from intact proviruses. As most defective proviruses are either hypermutated or contain large internal deletions (40), we analyzed 993 defective proviral sequences obtained from PLWH on ART and treated during chronic infection (https://psd.cancer.gov/). A total of 717 defective sequences (72.2%) would not be detected using our PCR approach. Hence, we conclude that our PCR and sequencing approach is likely to enrich for intact proviruses, although it is possible that these *env* sequences came from viruses with defects in other genes. Finally, all participants in this study were male, and whether these findings translate to women living with HIV requires further investigation.

Taken together, our results demonstrate a significant role for coreceptor tropism in HIV persistence on ART, with clear compartmentalization of infected T cell subsets on the basis of coreceptor usage, concordant with the coreceptor expression on those cells. These data are consistent with latency being established through direct infection of naive and central memory CD4$^+$ T cells by CXCR4-using strains. Identical sets of *env* sequences were found in multiple T cell subsets and subpopulations, consistent with infection of a common precursor and differentiation or dedifferentiation of infected cells. We therefore conclude that HIV persistence in different T cell subsets likely occurs through multiple mechanisms, including differentiation, proliferation, and *de novo* infection. This may potentially mean that different strategies will be needed to eliminate these infected cells.

MATERIALS AND METHODS

Cell lines. The embryonic kidney cell line 293T, JC53 cells (41), and TZM-bl cells (42) were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, 1% Glutamax, and 1% penicillin-streptomycin. The astrocytoma cell lines U87-CD4/CCR5 and U87-CD4/CXCR4 were maintained in DMEM supplemented with 15% fetal calf serum, 1% Glutamax, 1% penicillin-streptomycin, 1 μg/ml of puromycin, and 300 μg/ml of G418. Primary CD4$^+$ T cells were maintained in RPMI medium supplemented with 10% fetal calf serum, 1% Glutamax, and 1% penicillin-streptomycin.

Study participants and nucleic acid isolation. Samples were obtained from 8 HIV-infected participants enrolled in the SCOPE study at the University of California, San Francisco. These participants were treated during chronic infection; all had HIV RNA levels of $<$40 copies/ml and had been virally suppressed for $>$2 years. Participant characteristics, leukapheresis for obtaining blood CD4$^+$ T cells, sorting of CD4$^+$ T cell subsets, and biopsy of rectal tissue and inguinal lymph nodes for tissue-derived CD4$^+$ T cells have been described previously (12, 13). Briefly, using fluorescence-activated cell sorting, CD4$^+$ T cell subsets from blood and total CD4$^+$ T cells from tissue were isolated. The surface markers

used were as follows: naive ($CD4^+$ $CD45RA^+$ $CD27^+$ $CCR7^+$), central memory ($CD4^+$ $CD45RA^-$ $CD27^+$ $CCR7^+$), transitional memory ($CD4^+$ $CD45RA^-$ $CD27^+$ $CCR7^-$), effector memory ($CD4^+$ $CD45RA^-$ $CD27^-$ $CCR7^-$), and tissue total CD4 ($CD4^+$). Central memory cells were further sorted based on expression of CXCR3 and CCR6. Effector memory cells were further sorted based on expression of programmed cell death protein 1 (PD1). The subpopulations of CM and EM subsets were sorted because we and others have previously shown that the expression of PD1, CCR6, and CXCR3 on $CD4^+$ T cells from PLWH on ART has been associated with enrichment of HIV DNA (12, 13). In a subset of participants, $CD4^+$ T cells were sorted from rectal tissue ($n = 6$) and lymph node biopsy specimens ($n = 3$). Genomic DNA (gDNA) was extracted from 10^5 to 10^6 cells using the AllPrep DNA/RNA microkit (Qiagen) according to the manufacturer's protocol. Pretherapy plasma was obtained for two study participants 3 months (participant 2) and 29 months (participant 3) prior to the initiation of antiretrovirals (ARVs). Viral RNA (vRNA) from 1 ml of plasma was extracted using the QIAamp viral RNA minikit (Qiagen) according to the manufacturer's protocol. cDNA was synthesized using Superscript III (Invitrogen) and the subtype B-specific primer envB3out (5′-TTGCTACTTGTGATTGCTCCATGT). The level of integrated HIV DNA in each $CD4^+$ T cell subset was determined as previously described (2).

Single genome amplification for *env*. gDNA and cDNA were used as templates for single genome amplification of gp160 as previously described (43). Briefly, a nested PCR approach was performed using Platinum *Taq* HiFi (Invitrogen). The first-round primers were envB5out (5′-TAGAGCCCTGGAAGCATCCA GGAAG) and envB3out. One microliter of first-round PCR product was transferred to the second-round PCR using the primers envB5in (5′-TTAGGCATCTCCTATGGCAGGAAGAAG) and envB3in (5′-GTCTCGAGA TACTGCTCCCACCC). The second-round forward primer (envB5in) was further modified with the addition of CACC on the 5′ end to allow for TOPO directional cloning. To ensure single-copy input, the template nucleic acid was serially diluted to a point at which 2/8 nested gp160 PCRs were positive. The nested PCR was then performed again at the selected template dilution using a full 96-well plate. Positive PCRs were identified by gel electrophoresis, and the products were directly sequenced using Sanger sequencing.

Sequence analysis. A total of 722 *env* sequences were generated for this study (Table 2). *env* sequences were interrogated for the presence of hypermutation using the Hypermut2.0 tool (https://www.hiv.lanl.gov), and hypermutants were discarded from subsequent analysis. Inter- and intraindividual alignments of the V1 to V5 regions of *env* (1,019 bp) were performed using Clustal W in CLC Main Workbench (Qiagen). For phylogenetic analysis, maximum likelihood trees were constructed using MEGA 6 (http://megasoftware.net). The general time reversible plus gamma model of nucleotide substitution was used. This model was selected by using the Los Alamos National Laboratory FindModel program (https://www.hiv.lanl.gov). The phylogenetic tree structures were statistically supported by 100 bootstrap replicates. Identical sequences were identified by calculating average pairwise distances in MEGA 6 and confirmed with individual sequence alignments. Sequences differing by 1 bp were considered unique. Our sequencing approach has a clonal predictive score of 83% (44). Diversity within subsets was determined by calculating average pairwise distances in MEGA 7 using the maximum composite likelihood model of nucleotide substitution (transitions and transversions). Variance estimation was calculated using 1,000 bootstrap replicates. Genetic compartmentalization was determined by Panmixia tests using 10,000 permutations as previously described (45). Diversity and compartmentalization tests were performed in subsets with a minimum of 4 nonhypermutated *env* sequences. No contamination was observed between participants. Coreceptor tropism was determined using the geno2pheno coreceptor prediction algorithm with a 5% false-positive rate (FPR) cutoff (16).

***env* cloning.** Selected *env* PCR amplimers were cloned into the pcDNA3.1 Directional TOPO expression vector (Thermo Fisher Scientific) according to the manufacturer's protocol. Clones containing the correct insert were screened using a colony PCR approach.

Production and titration of Env-pseudotyped luciferase and GFP reporter viruses. Env-pseudotyped luciferase or GFP reporter viruses were produced by transfection of 293T cells with plasmids pcDNA3.1-Env and pNL4.3e-luc or pNL4.3e-GFP using Polyethylenimine (PEI) Max (Polysciences) at a ratio of 1:4. Supernatants containing luciferase reporter viruses were harvested 48 h posttransfection, filtered through 0.45-μm-pore-size filters, aliquoted, and stored at −80°C. Supernatants containing GFP reporter viruses were harvested 48 h posttransfection, filtered through 0.45-μm-pore-size filters, and concentrated over a 25% (wt/vol) sucrose cushion via ultracentrifugation at 25,000 rpm for 2 h at 4°C. The 50% tissue culture infectious dose ($TCID_{50}$) was determined by titration in JC53 cells (41) for luciferase reporter viruses and TZM-bl cells for GFP reporter viruses as described previously (46, 47).

Coreceptor tropism assays. Env coreceptor tropism was determined by infection of U87-CD4/CCR5 and U87-CD4/CXCR4 with Env-pseudotyped luciferase reporter viruses as previously described (48). Briefly, 1×10^4 U87-CD4/CCR5 or U87-CD4/CXCR4 cells were seeded per well into 96-well flat-bottom plates. The following day, the cells were inoculated with 5-fold serial dilutions (neat to 1/125) of Env-pseudotyped luciferase reporter viruses in a volume of 200 μl for 12 h. The viral inoculum was removed and replaced with fresh medium, and the cells were incubated for a further 60 h. The level of reporter virus entry was measured by luciferase activity in cell lysates (Promega) according to the manufacturer's instructions. Luminescence was measured using a FLUOStar microplate reader (BMG Labtech). Background luciferase activity was determined using mock-infected cells that were inoculated with medium only and by infections of U87-CD4/CCR5 cells with reporter viruses pseudotyped with the X4-tropic Env NL4.3 and infections of U87-CD4/CXCR4 cells with reporter viruses pseudotyped with the R5-tropic Env AD8.

Maraviroc sensitivity assays. Sensitivity to maraviroc (MVC) was determined by infection of U87-CD4/CCR5 cells incubated with serial dilutions of MVC as previously described (49). Briefly, 1×10^4 U87-CD4/CCR5 cells were seeded per well into 96-well flat-bottom plates. The following day, the cells

were preincubated with 5-fold serial dilutions of MVC (1,000 nM to 0.00256 nM) for a period of 30 min prior to the addition of 100 $TCID_{50}$s of Env-pseudotyped luciferase reporter viruses. The virus-containing inoculum was washed off after 12 h, replaced with fresh medium, and further incubated for a total period of 72 h. MVC was maintained throughout the experiment, and the dimethyl sulfoxide (DMSO) concentration was consistent across all wells (0.1%). The level of reporter virus entry was determined as described above, and virus entry in wells containing MVC was normalized to that of wells incubated with DMSO. Inhibition curves were generated with a nonlinear function and least-squares analysis in Prism (GraphPad). The IC_{50} and IC_{90} were determined from these curves.

T cell tropism assays. Peripheral blood mononuclear cells (PBMC) were isolated from buffy coats of HIV-negative healthy donors. Resting $CD4^+$ T cells were purified using the EasySep human $CD4^+$ T cell isolation kit (StemCell Technologies) according to the manufacturer's instructions. For each experiment, >98% of the cells were $CD3^+$, $CD4^+$, and $CD69^-$. Cells were either incubated in medium alone or incubated in medium supplemented with 20 μg/ml of phytohemagglutinin (PHA) and 10 U/ml of interleukin 2 (IL-2) for 48 h at 37°C. Subsequent to this, 5×10^5 resting or activated $CD4^+$ T cells were infected with equivalent amounts of Env-pseudotyped GFP reporter viruses using spinoculation (1,200 $\times$ *g* for 2 h at room temperature) in 100 μl. The viral inoculum was then removed, and cells were maintained in RF10 supplemented with 1 U/ml of IL-2 for 5 days. Cells were then stained with a LIVE/DEAD stain (Aqua DEAD cell stain kit; Life Technologies) and fixed with 1% paraformaldehyde. For flow cytometry analysis, <100,000 events were collected on an LSR Fortessa flow cytometer (BD Biosciences) and analyzed with FlowJo software (FlowJo LLC).

Data availability. The accession numbers for all sequences reported in this study are GenBank numbers MK465705 to MK466337.

ACKNOWLEDGMENTS

We thank D. Kabat for providing JC53 cells. U87-CD4/CCR5 cells, U87-CD4/CXCR4 cells, TZM-bl cells, maraviroc, and AMD3100 were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH.

S.G.D. received grant support from Gilead, Merck, and ViiV, has consulted for Janssen and Abbvie, and is on the scientific advisory boards for Enochian Biosciences and Bryologix. S.R.L. has received investigator-initiated grant funding from Gilead, Merck, and ViiV. She has provided paid scientific advice to Abivax, Abbvie, and Gilead. The other authors declare no conflicts of interest.

REFERENCES

1. Deeks SG, Lewin SR, Ross AL, Ananworanich J, Benkirane M, Cannon P, Chomont N, Douek D, Lifson JD, Lo YR, Kuritzkes D, Margolis D, Mellors J, Persaud D, Tucker JD, Barre-Sinoussi F, International AIDS Society Towards a Cure Working Group, Alter G, Auerbach J, Autran B, Barouch DH, Behrens G, Cavazzana M, Chen Z, Cohen EA, Corbelli GM, Eholie S, Eyal N, Fidler S, Garcia L, Grossman C, Henderson G, Henrich TJ, Jefferys R, Kiem HP, McCune J, Moodley K, Newman PA, Nijhuis M, Nsubuga MS, Ott M, Palmer S, Richman D, Saez-Cirion A, Sharp M, Siliciano J, Silvestri G, Singh J, Spire B, Taylor J, Tolstrup M, Valente S, van Lunzen J, Walensky R, Wilson I, Zack J. 2016. International AIDS Society global scientific strategy: towards an HIV cure 2016. Nat Med 22:839–850. https://doi.org/10.1038/nm.4108.
2. 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.
3. Wightman F, Solomon A, Khoury G, Green JA, Gray L, Gorry PR, Ho YS, Saksena NK, Hoy J, Crowe SM, Cameron PU, Lewin SR. 2010. Both CD31(+) and CD31(−) naive CD4(+) T cells are persistent HIV type 1-infected reservoirs in individuals receiving antiretroviral therapy. J Infect Dis 202:1738–1748. https://doi.org/10.1086/656721.
4. Hosmane NN, Kwon KJ, Bruner KM, Capoferri AA, Beg S, Rosenbloom DI, Keele BF, Ho YC, Siliciano JD, Siliciano RF. 2017. Proliferation of latently infected CD4+ T cells carrying replication-competent HIV-1: potential role in latent reservoir dynamics. J Exp Med 214:959–972. https://doi.org/10.1084/jem.20170193.
5. 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.
6. Bui JK, Sobolewski MD, Keele BF, Spindler J, Musick A, Wiegand A, Luke BT, Shao W, Hughes SH, Coffin JM, Kearney MF, Mellors JW. 2017. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. PLoS Pathog 13:e1006283. https://doi.org/10.1371/journal.ppat.1006283.
7. Gorry PR, Ancuta P. 2011. Coreceptors and HIV-1 pathogenesis. Curr HIV/AIDS Rep 8:45–53. https://doi.org/10.1007/s11904-010-0069-x.
8. Bleul CC, Wu L, Hoxie JA, Springer TA, Mackay CR. 1997. The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. Proc Natl Acad Sci U S A 94:1925–1930. https://doi.org/10.1073/pnas.94.5.1925.
9. Flynn JK, Paukovics G, Cashin K, Borm K, Ellett A, Roche M, Jakobsen MR, Churchill MJ, Gorry PR. 2014. Quantifying susceptibility of CD4+ stem memory T-cells to infection by laboratory adapted and clinical HIV-1 strains. Viruses 6:709–726. https://doi.org/10.3390/v6020709.
10. Flynn JK, Paukovics G, Moore MS, Ellett A, Gray LR, Duncan R, Salimi H, Jubb B, Westby M, Purcell DF, Lewin SR, Lee B, Churchill MJ, Gorry PR, Roche M. 2013. The magnitude of HIV-1 resistance to the CCR5 antagonist maraviroc may impart a differential alteration in HIV-1 tropism for macrophages and T-cell subsets. Virology 442:51–58. https://doi.org/10.1016/j.virol.2013.03.026.
11. Pfaff JM, Wilen CB, Harrison JE, Demarest JF, Lee B, Doms RW, Tilton JC. 2010. HIV-1 resistance to CCR5 antagonists associated with highly efficient use of CCR5 and altered tropism on primary CD4+ T cells. J Virol 84:6505–6514. https://doi.org/10.1128/JVI.00374-10.
12. Fromentin R, Bakeman W, Lawani MB, Khoury G, Hartogensis W, DaFonseca S, Killian M, Epling L, Hoh R, Sinclair E, Hecht FM, Bacchetti P, Deeks SG, Lewin SR, Sekaly RP, Chomont N. 2016. CD4+ T cells expressing PD-1, TIGIT and LAG-3 contribute to HIV persistence during ART. PLoS Pathog 12:e1005761. https://doi.org/10.1371/journal.ppat.1005761.
13. Khoury G, Anderson JL, Fromentin R, Hartogenesis W, Smith MZ, Bacchetti P, Hecht FM, Chomont N, Cameron PU, Deeks SG, Lewin SR. 2016. Persistence of integrated HIV DNA in CXCR3+ CCR6+ memory CD4+ T cells in HIV-infected individuals on antiretroviral therapy. AIDS 30:1511–1520. https://doi.org/10.1097/QAD.0000000000001029.
14. Khoury G, Fromentin R, Solomon A, Hartogensis W, Killian M, Hoh R,

Somsouk M, Hunt PW, Girling V, Sinclair E, Bacchetti P, Anderson JL, Hecht FM, Deeks SG, Cameron PU, Chomont N, Lewin SR. 2017. Human immunodeficiency virus persistence and T-cell activation in blood, rectal, and lymph node tissue in human immunodeficiency virus-infected individuals receiving suppressive antiretroviral therapy. J Infect Dis 215:911–919. https://doi.org/10.1093/infdis/jix039.

15. Estes JD, Kityo C, Ssali F, Swainson L, Makamdop KN, Del Prete GQ, Deeks SG, Luciw PA, Chipman JG, Beilman GJ, Hoskuldsson T, Khoruts A, Anderson J, Deleage C, Jasurda J, Schmidt TE, Hafertepe M, Callisto SP, Pearson H, Reimann T, Schuster J, Schoephoerster J, Southern P, Perkey K, Shang L, Wietgrefe SW, Fletcher CV, Lifson JD, Douek DC, McCune JM, Haase AT, Schacker TW. 2017. Defining total-body AIDS-virus burden with implications for curative strategies. Nat Med 23:1271–1276. https://doi.org/10.1038/nm.4411.
16. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. 2007. Bioinformatics prediction of HIV coreceptor usage. Nat Biotechnol 25:1407–1410. https://doi.org/10.1038/nbt1371.
17. Parrish NF, Wilen CB, Banks LB, Iyer SS, Pfaff JM, Salazar-Gonzalez JF, Salazar MG, Decker JM, Parrish EH, Berg A, Hopper J, Hora B, Kumar A, Mahlokozera T, Yuan S, Coleman C, Vermeulen M, Ding H, Ochsenbauer C, Tilton JC, Permar SR, Kappes JC, Betts MR, Busch MP, Gao F, Montefiori D, Haynes BF, Shaw GM, Hahn BH, Doms RW. 2012. Transmitted/founder and chronic subtype C HIV-1 use CD4 and CCR5 receptors with equal efficiency and are not inhibited by blocking the integrin alpha4beta7. PLoS Pathog 8:e1002686. https://doi.org/10.1371/journal.ppat.1002686.
18. Sterjovski J, Roche M, Churchill MJ, Ellett A, Farrugia W, Gray LR, Cowley D, Poumbourios P, Lee B, Wesselingh SL, Cunningham AL, Ramsland PA, Gorry PR. 2010. An altered and more efficient mechanism of CCR5 engagement contributes to macrophage tropism of CCR5-using HIV-1 envelopes. Virology 404:269–278. https://doi.org/10.1016/j.virol.2010.05.006.
19. De Clercq E, Yamamoto N, Pauwels R, Balzarini J, Witvrouw M, De Vreese K, Debyser Z, Rosenwirth B, Peichl P, Datema R. 1994. Highly potent and selective inhibition of human immunodeficiency virus by the bicyclam derivative JM3100. Antimicrob Agents Chemother 38:668–674. https://doi.org/10.1128/aac.38.4.668.
20. 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.
21. 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.
22. Cheret A, Bacchus-Souffan C, Avettand-Fenoel V, Melard A, Nembot G, Blanc C, Samri A, Saez-Cirion A, Hocqueloux L, Lascoux-Combe C, Allavena C, Goujard C, Valantin MA, Leplatois A, Meyer L, Rouzioux C, Autran B, OPTIPRIM ANRS-147 Study Group. 2015. Combined ART started during acute HIV infection protects central memory CD4+ T cells and can induce remission. J Antimicrob Chemother 70:2108–2120.
23. Saez-Cirion A, Bacchus C, Hocqueloux L, Avettand-Fenoel V, Girault I, Lecuroux C, Potard V, Versmisse P, Melard A, Prazuck T, Descours B, Guergnon J, Viard JP, Boufassa F, Lambotte O, Goujard C, Meyer L, Costagliola D, Venet A, Pancino G, Autran B, Rouzioux C, ANRS VISCONTI Study Group. 2013. Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. PLoS Pathog 9:e1003211. https://doi.org/10.1371/journal.ppat.1003211.
24. Heeregrave EJ, Geels MJ, Brenchley JM, Baan E, Ambrozak DR, van der Sluis RM, Bennemeer R, Douek DC, Goudsmit J, Pollakis G, Koup RA, Paxton WA. 2009. Lack of in vivo compartmentalization among HIV-1 infected naive and memory CD4+ T cell subsets. Virology 393:24–32. https://doi.org/10.1016/j.virol.2009.07.011.
25. Ostrowski MA, Chun TW, Justement SJ, Motola I, Spinelli MA, Adelsberger J, Ehler LA, Mizell SB, Hallahan CW, Fauci AS. 1999. Both memory and CD45RA$^+$/CD62L$^+$ naive CD4$^+$ T cells are infected in human immunodeficiency virus type 1-infected individuals. J Virol 73:6430–6435. https://doi.org/10.1128/JVI.73.8.6430-6435.1999.
26. Blaak H, van't Wout AB, Brouwer M, Hooibrink B, Hovenkamp E, Schuitemaker H. 2000. In vivo HIV-1 infection of CD45RA(+)CD4(+) T cells is established primarily by syncytium-inducing variants and correlates with the rate of CD4(+) T cell decline. Proc Natl Acad Sci U S A 97:1269–1274. https://doi.org/10.1073/pnas.97.3.1269.
27. Puertas MC, Noguera-Julian M, Massanella M, Pou C, Buzon MJ, Clotet B, Stevenson M, Paredes R, Blanco J, Martinez-Picado J. 2016. Lack of concordance between residual viremia and viral variants driving de novo infection of CD4(+) T cells on ART. Retrovirology 13:51. https://doi.org/10.1186/s12977-016-0282-9.
28. Kwa D, Vingerhoed J, Boeser-Nunnink B, Broersen S, Schuitemaker H. 2001. Cytopathic effects of non-syncytium-inducing and syncytium-inducing human immunodeficiency virus type 1 variants on different CD4(+)-T-cell subsets are determined only by coreceptor expression. J Virol 75:10455–10459. https://doi.org/10.1128/JVI.75.21.10455-10459.2001.
29. Pierson T, Hoffman TL, Blankson J, Finzi D, Chadwick K, Margolick JB, Buck C, Siliciano JD, Doms RW, Siliciano RF. 2000. Characterization of chemokine receptor utilization of viruses in the latent reservoir for human immunodeficiency virus type 1. J Virol 74:7824–7833. https://doi.org/10.1128/jvi.74.17.7824-7833.2000.
30. Cashin K, Paukovics G, Jakobsen MR, Ostergaard L, Churchill MJ, Gorry PR, Flynn JK. 2014. Differences in coreceptor specificity contribute to alternative tropism of HIV-1 subtype C for CD4(+) T-cell subsets, including stem cell memory T-cells. Retrovirology 11:97. https://doi.org/10.1186/s12977-014-0097-5.
31. 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 reprogramming 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.
32. Paiardini M, Cervasi B, Reyes-Aviles E, Micci L, Ortiz AM, Chahroudi A, Vinton C, Gordon SN, Bosinger SE, Francella N, Hallberg PL, Cramer E, Schlub T, Chan ML, Riddick NE, Collman RG, Apetrei C, Pandrea I, Else J, Munch J, Kirchhoff F, Davenport MP, Brenchley JM, Silvestri G. 2011. Low levels of SIV infection in sooty mangabey central memory CD(4)(+) T cells are associated with limited CCR5 expression. Nat Med 17:830–836. https://doi.org/10.1038/nm.2395.
33. Siliciano RF, Greene WC. 2011. HIV latency. Cold Spring Harb Perspect Med 1:a007096. https://doi.org/10.1101/cshperspect.a007096.
34. Youngblood B, Hale JS, Kissick HT, Ahn E, Xu X, Wieland A, Araki K, West EE, Ghoneim HE, Fan Y, Dogra P, Davis CW, Konieczny BT, Antia R, Cheng X, Ahmed R. 2017. Effector CD8 T cells dedifferentiate into long-lived memory cells. Nature 552:404–409. https://doi.org/10.1038/nature25144.
35. Evans VA, Kumar N, Filali A, Procopio FA, Yegorov O, Goulet JP, Saleh S, Haddad EK, da Fonseca Pereira C, Ellenberg PC, Sekaly RP, Cameron PU, Lewin SR. 2013. Myeloid dendritic cells induce HIV-1 latency in non-proliferating CD4+ T cells. PLoS Pathog 9:e1003799. https://doi.org/10.1371/journal.ppat.1003799.
36. Kumar NA, van der Sluis RM, Mota T, Pascoe R, Evans VA, Lewin SR, Cameron PU. 2018. Myeloid dendritic cells induce HIV latency in proliferating CD4(+) T cells. J Immunol 201:1468–1477. https://doi.org/10.4049/jimmunol.1701233.
37. 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.
38. Boritz EA, Darko S, Swaszek L, Wolf G, Wells D, Wu X, Henry AR, Laboune F, Hu J, Ambrozak D, Hughes MS, Hoh R, Casazza JP, Vostal A, Bunis D, Nganou-Makamdop K, Lee JS, Migueles SA, Koup RA, Connors M, Moir S, Schacker T, Maldarelli F, Hughes SH, Deeks SG, Douek DC. 2016. Multiple origins of virus persistence during natural control of HIV infection. Cell 166:1004–1015. https://doi.org/10.1016/j.cell.2016.06.039.
39. Hiener B, Horsburgh BA, Eden JS, 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.
40. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent non-

induced 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.
41. Platt EJ, Wehrly K, Kuhmann SE, Chesebro B, Kabat D. 1998. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. J Virol 72:2855–2864. https://doi.org/10.1128/JVI.72.4.2855-2864.1998.
42. Wei X, Decker JM, Liu H, Zhang Z, Arani RB, Kilby JM, Saag MS, Wu X, Shaw GM, Kappes JC. 2002. Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. Antimicrob Agents Chemother 46:1896–1905. https://doi.org/10.1128/aac.46.6.1896-1905.2002.
43. Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, Keele BF, Derdeyn CA, Farmer P, Hunter E, Allen S, Manigart O, Mulenga J, Anderson JA, Swanstrom R, Haynes BF, Athreya GS, Korber BT, Sharp PM, Shaw GM, Hahn BH. 2008. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. J Virol 82:3952–3970. https://doi.org/10.1128/JVI.02660-07.
44. 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.
45. Achaz G, Palmer S, Kearney M, Maldarelli F, Mellors JW, Coffin JM, Wakeley J. 2004. A robust measure of HIV-1 population turnover within chronically infected individuals. Mol Biol Evol 21:1902–1912. https://doi.org/10.1093/molbev/msh196.
46. Gray L, Roche M, Churchill MJ, Sterjovski J, Ellett A, Poumbourios P, Sherieff S, Sheffief S, Wang B, Saksena N, Purcell DFJ, Wesselingh S, Cunningham AL, Brew BJ, Gabuzda D, Gorry PR. 2009. Tissue-specific sequence alterations in the human immunodeficiency virus type 1 envelope favoring CCR5 usage contribute to persistence of dual-tropic virus in the brain. J Virol 83:5430–5441. https://doi.org/10.1128/JVI.02648-08.
47. Sterjovski J, Churchill MJ, Ellett A, Gray LR, Roche MJ, Dunfee RL, Purcell DF, Saksena N, Wang B, Sonza S, Wesselingh SL, Karlsson I, Fenyo EM, Gabuzda D, Cunningham AL, Gorry PR. 2007. Asn 362 in gp120 contributes to enhanced fusogenicity by CCR5-restricted HIV-1 envelope glycoprotein variants from patients with AIDS. Retrovirology 4:89. https://doi.org/10.1186/1742-4690-4-89.
48. Jakobsen MR, Cashin K, Roche M, Sterjovski J, Ellett A, Borm K, Flynn J, Erikstrup C, Gouillou M, Gray LR, Saksena NK, Wang B, Purcell DF, Kallestrup P, Zinyama-Gutsire R, Gomo E, Ullum H, Ostergaard L, Lee B, Ramsland PA, Churchill MJ, Gorry PR. 2013. Longitudinal analysis of CCR5 and CXCR4 usage in a cohort of antiretroviral therapy-naive subjects with progressive HIV-1 subtype C infection. PLoS One 8:e65950. https://doi.org/10.1371/journal.pone.0065950.
49. Roche M, Jakobsen MR, Ellett A, Salimiseyedabad H, Jubb B, Westby M, Lee B, Lewin SR, Churchill MJ, Gorry PR. 2011. HIV-1 predisposed to acquiring resistance to maraviroc (MVC) and other CCR5 antagonists in vitro has an inherent, low-level ability to utilize MVC-bound CCR5 for entry. Retrovirology 8:89. https://doi.org/10.1186/1742-4690-8-89.