Published in final edited form as:

AIDS Res Hum Retroviruses. 2018 July ; 34(7): 607–620. doi:10.1089/AID.2017.0209.

Individuals with HIV-1 subtype C infection and cryptococcal meningitis exhibit viral genetic intermixing of HIV-1 between plasma and cerebrospinal fluid, and a high prevalence of CXCR4-using variants

Katlego Sojane¹, Richard T. Kangethe¹, Christina C. Chang², Mahomed-Yunus S. Moosa³, Sharon R. Lewin^{2,4}, Martyn A. French^{5,6}, and Thumbi Ndung'u^{1,7,8,9}

¹HIV Pathogenesis Programme, Doris Duke Medical Research Institute, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa

²Department of Infectious Diseases, Alfred Hospital and Monash University, Melbourne, Australia

³Department of Infectious Diseases, King Edward VIII Hospital, University of KwaZulu-Natal, Durban, South Africa

⁴The Peter Doherty Institute for Infection and Immunity, University of Melbourne and Royal Melbourne Hospital, Melbourne, Australia

⁵Medical School and School of Biomedical Sciences, University of Western Australia, Perth, Australia

⁶Department of Clinical Immunology, Royal Perth Hospital and PathWest Laboratory Medicine, Perth, Australia

⁷Africa Health Research Institute, Durban, South Africa

⁸Ragon Institute of MGH, MIT and Harvard University, Cambridge, MA 02139, USA

⁹Max Planck Institute for Infection Biology, Berlin, Germany

Abstract

The genotypic properties of HIV-1 subtype C in individuals presenting with cryptococcal meningitis (CM) are not well established. Employing single-genome amplification as well as bulk PCR, cloning and sequencing strategies, we evaluated the genetic properties of HIV-1 subtype C *env* in 16 antiretroviral therapy naïve study participants with CM. Eleven of the 16 participants had matched blood plasma and cerebrospinal fluid (CSF) evaluated, with the rest having either a plasma or CSF sample evaluated. Before antiretroviral therapy initiation, matched plasma and cerebrospinal fluid-derived *env* sequences of all 11 participants displayed genetic intermixing

Address correspondence to: Thumbi Ndung'u, HIV Pathogenesis Programme, Doris Duke Medical Research Institute, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, 719 Umbilo Road, Congella, Durban, 4001. ndungu@ukzn.ac.za.

Sequence data

Nucleotide sequences are available in GenBank with the following accession numbers: MF284812-MF284921 and MF284922-MF285062.

Competing interests

The authors have no competing interests.

between the two compartments. Overall, seven of the 16 (~43.8%) participants harbored CXCR4-using variants in plasma and/or CSF, according to coreceptor usage prediction algorithms. This study suggests that HIV-1 subtype C genetic intermixing between peripheral blood and the central nervous system is common in individuals presenting with CM, and that CXCR4 usage is present in one or both compartments in approximately 44% of individuals.

Keywords

Cerebrospinal fluid; cryptococcal meningitis; HIV-1; plasma

Introduction

To initiate infection, the human immunodeficiency virus type 1 (HIV-1) attaches itself to a target cell by binding to cell-surface CD4 and a chemokine receptor, usually CCR5 or CXCR4.^{1–3} The ability of HIV-1 to use a particular coreceptor for entry is determined primarily by the properties of the envelope protein (Env) V3 loop^{4, 5}, though determinants outside of the V3 loop may also contribute.^{6–9} Variants which use CCR5 only (R5-tropic) predominate early in infection generally, however, variants with expanded coreceptor usage ability may arise over time.^{10, 11} Approximately 50% of individuals infected with HIV-1 subtype B (HIV-1B) have variants which use CCR5 and CXCR4 (R5X4-tropic) or CXCR4 only (X4-tropic) during the chronic stage of disease, and the emergence of these variants is associated with an accelerated decline of CD4⁺ T-cell count.^{12, 13} In contrast, the prevalence of CXCR4 usage for entry by HIV-1 subtype C (HIV-1C) is unclear. Some studies have reported an absence or low prevalence of CXCR4 usage by HIV-1C^{14–18}, while others have detected 15% to 30% of individuals with advanced or end-stage disease harboring CXCR4-using variants.^{19–21} Further studies are needed to clarify the frequency and circumstances of CXCR4 usage by HIV-1C isolates, so that individuals' eligibility for use of coreceptor blocking drugs such as maraviroc, a CCR5 antagonist, may be accurately determined.

HIV-1 is detectable in the central nervous system (CNS) in primary infection^{22–24}, where its presence may ultimately cause HIV-associated neurological disorders.^{25, 26} Studies of CNS compartmentalization of HIV-1 are limited because obtaining brain biopsies is not possible, however, the cerebrospinal fluid (CSF) has been used often as a surrogate of the CNS compartment.^{27–29} Early studies of HIV-1B *env* genes from matched plasma and CSF showed that variants unique to the CSF could arise³⁰, and that this genetic compartmentalization of variants was most pronounced in individuals with HIV-associated dementia (HAD).^{31, 32} Later studies combining genotypic and/or phenotypic analyses confirmed clonal amplification of HIV-1B and HIV-1C variants in the CNS, indicating that the CNS can act as a compartment of unique HIV persistence relative to peripheral blood.^{24, 33, 34} In a limited number of studies, a discordant entry phenotype of HIV-1B between the blood and CNS was reported in individuals with chronic stage HIV-1 disease.^{27, 35–37} It is unclear how common genotypic discordance of HIV-1C between plasma and CSF is in individuals with advanced HIV infection and a co-infection.

Individuals harboring R5 HIV-1 only are eligible for combination antiretroviral therapy (cART) containing maraviroc, a CCR5 antagonist, while those harboring R5X4 or X4-tropic viruses are precluded from using this drug.³⁸ There are multiple genotypic tools or algorithms used to determine the coreceptor usage of HIV-1 in order to determine eligibility for use of maraviroc.³⁹ One approach is to use coreceptor usage prediction algorithms (CPAs) such as Geno2pheno, WebPSSM and PhenoSeq which utilize the whole V3 loop sequences of HIV-1 Envs to predict whether a virus will be R5 or CXCR4-using (using CCR5 and CXCR4, or CXCR4 alone).^{40–42} CPAs, therefore, are attractive diagnostic tools, as they are relatively inexpensive and less laborious to use than determinations by *in vitro* phenotypic assays.

HIV-related cryptococcal meningitis (CM) is an acquired immune deficiency syndrome (AIDS) defining illness which is associated with high acute mortality rates, and is most prevalent in sub-Saharan Africa.^{43–45} Previous studies by our group revealed that compartmentalization of immune responses may arise between the blood and CNS compartment in individuals with CM, where pro-inflammatory natural killer cells, non-classical monocytes, chemokines and cytokines are enriched in CSF relative to blood.⁴⁶ We were interested in establishing whether or not HIV-1 genetic properties are compartmentalized in individuals with HIV-related CM. We hypothesized that localized inflammation in the CNS caused by *Cryptococcus neoformans* (or its antigens) provides an immunological milieu that recruits HIV-1C variants into the CNS, causing viral genetic mixing between peripheral blood and the CNS. We therefore explored whether CM was associated with HIV-1C genetic compartmentalization between blood plasma and the CSF or not, and whether in this setting there was evidence of CXCR4 usage by HIV-1 subtype C variants.

Materials and methods

Study participants

The participants for this sub-study were part of a larger prospective clinical study of Cryptococcosis-associated immune reconstitution inflammatory syndrome (C-IRIS) conducted in Durban, South Africa between August 2009 and September 2011, as described in detail elsewhere.⁴⁷ Written informed consent was provided by each participant or their next-of-kin. The ethics review boards of the University of KwaZulu-Natal (reference number BF053/09), Monash University (2009001224) and University of Western Australia (RA/4/1/2541) granted ethics approval for the study. For the current study, 16 participants infected with HIV-1 and presenting with CM for the first time were randomly selected based on the criteria of detectable HIV-1 RNA in the plasma and CSF according to the COBAS TaqMan HIV-1 test (Hoffmann-La Roche, Basel, Switzerland) with a lower limit of 34 copies/mL, a CD4⁺ T-cell count less than 200 cells/ μ L, and either a positive CSF cryptococcal antigen (CrAg) or Indian ink test result.

Plasma and CSF sample processing

The cell-free HIV-1 RNA levels in matched plasma and CSF samples were measured in the same polymerase chain reaction (PCR) assay run using the COBAS TaqMan HIV-1 test as

described previously.⁴⁸ Routine cell count analyses including the enumeration of white blood cell subsets in the CSF and CD4⁺ T-cells in peripheral blood were performed by standard clinical laboratory assays in an accredited laboratory.⁴⁷

RNA extraction and the synthesis of cDNA

The extraction of HIV-1 RNA from the plasma and CSF was performed using the QIAamp Viral RNA Mini kit (Qiagen, Dusseldorf, Germany) according to the manufacturer's instructions. SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA) was used for cDNA synthesis as previously described.⁴⁹ Briefly, 10 µL of extracted RNA was added to a mixture of 1 mM of deoxynucleoside triphosphates and 2 µM of primer OFM19 (5'-GCACTCAAGGCAAGCTTTATTGAGGCTTA-3'; HXB2 positions 9,604-9,632). The mixture was first heated to 65 °C for five minutes and cooled to 4 °C for one minute. The following reagents were combined and then added to the initial mixture for a final volume of 20 µL per sample: 1 X reverse transcriptase (RT) buffer, 0.1 M dithiothreitol, 40 units of RNaseOUT (Invitrogen), and 200 units of SuperScript III Reverse Transcriptase enzyme. The complete reaction mixture was incubated at 50 °C for 60 minutes, 55 °C for 60 minutes to commence cDNA synthesis, incubated at 70 °C for 15 minutes to inactivate the reverse transcriptase enzyme, and cooled to 4 °C to end cDNA synthesis.

Single genome amplification of plasma and CSF-derived variants

Unique viral variants were purified by single-genome amplification (SGA) and Sanger sequencing as described previously.⁴⁹ Synthesized cDNA was serially diluted, and a minimum of 10 PCRs for each dilution were performed. The dilution which presented less than or equal to 30% positive amplification reactions represented the end-point dilution, where each positive amplification reaction contained a unique viral template.⁵⁰ Both rounds of PCR were performed using the Platinum Taq DNA Polymerase High Fidelity (Invitrogen) as described previously.^{49, 51} The first round of PCR included the forward primer VIF1 (5'-GGGTTTATTACAGGGACAGCAGAG-3'; HXB2 positions 4,900-4,923), reverse primer OFM19, and 1 µL of diluted cDNA. The second round of PCR included forward primer ENVA (5'-GCTTAGGCATCTCCTATGGCAGGAAGAA-3'; HXB2 positions 5,945-5,982), reverse primer ENVN (5'-CTGCCAATCAGGGAAGTAGCCTTGTGT-3'; HXB2 positions 9,145-9,171) and 2 µL of first round PCR product. The cycling conditions for both rounds of PCR were as follows: 94 °C for 4 minutes; 35 cycles of 94 °C for 15 seconds, 55 °C for 30 seconds, 68 °C for 4 minutes; a final extension step of 68 °C step for 20 minutes followed by a hold at 4 °C. Positive PCR reactions were detected by agarose gel electrophoresis, and amplicons from end-point dilution reactions were sequenced directly from the second round PCR products.

Bulk PCR and cloning

Using a modified, nested bulk PCR and cloning protocol, 2.1 kb KpnI-to-BamHI *env* gene fragments were amplified from undiluted cDNA as described previously.^{52–54} Briefly, the forward VIF1 primer, reverse OFM19 primer, the Platinum Taq DNA Polymerase High Fidelity (Invitrogen), and 1 µL of undiluted cDNA were included in the first round PCR reaction mixture. In the second round of PCR, the forward Env-KpnI primer (5'-GTCTATTATGGGGTACCTGTGTGG-3'; HXB2 positions 6,336-6,359), reverse Env-

BamHI primer (5'-GCTAAGGATCCGTTCTACTAATCGT-3'; HXB2 position 8,463-8,485), the Phusion High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA), and 2 μ L of PCR product from the first round of PCR were included. PCR products from second round reactions were analyzed by agarose gel electrophoresis, and detected amplicons were gel purified. Purified amplicons were cloned into the pSVIII-Env plasmid for the purposes of sequencing.^{52, 54, 55}

Sequencing of HIV-1 *env* genes

SGA-derived amplicons from second round PCR products and recombinant pSVIII-Env vectors were sequenced using the ABI Prism Big Dye Terminator Version 3.1 cycle sequencing kit (Applied Biosystems, Waltham, MA), incorporating multiple forward and reverse primers that covered the entire HIV-1 *env* gene.⁵⁶ Sequences were resolved using the ABI 3130 XL genetic analyzer and contigs composing *env* genes were assembled and manually edited with Sequencher version 5.4.1 (Genecodes, Ann Arbor, MI). Sequencing chromatograms with multiple overlapping peaks were discarded along with duplicate sequences.

Phylogenetic analyses

All participant-derived *env* genes were aligned with a panel of reference sequences from different subtypes available in the Los Alamos National Laboratory HIV database (<https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html>) using Clustal W. With MEGA 7 software⁵⁷, a neighbor-joining phylogenetic tree of 1,000 bootstrap replicates was constructed using the V1 to V5 *env* gene regions to determine the subtype, detect inter-participant sequence contamination, and identify inter-compartment mixing of HIV-1 variants. Evolutionary distances were computed using the maximum-likelihood composite method, where all positions containing gaps or missing data were excluded in the analysis as described previously.¹⁵

Compartmentalization analyses

For two participants, CM112 and CM089, there were a minimum of 10 SGA-derived sequences from the plasma and CSF each, allowing for evaluation of the degree of HIV-1 genetic compartmentalization between the two compartments as described previously.^{33, 58} Maximum-likelihood phylogenetic trees showing the intra-participant phylogenetic relatedness of plasma and CSF-derived *env* genes were generated using MEGA 7⁵⁷ and PhyML software⁵⁹ with the following parameters: HK85 nucleotide substitution model, four substitution rate categories, a transition/transversion ratio of four, estimated gamma distribution parameter and 1,000 bootstrap replicates. Compartmentalization of CSF sequences was determined using the Slatkin-Maddison (tree-based) and Hudson nearest-neighbor tests (distance-based) with 10,000 permutations, in HYPHY as previously described.^{33, 58, 60} A *p*-value equal to or lower than 0.05 signified CSF compartmentalization of sequences.

Prediction of coreceptor usage

The publicly available coreceptor usage prediction algorithms (CPAs), Geno2pheno with a false positive rate (FPR) of 5% (<http://coreceptor.bioinf.mpi-inf.mpg.de/>), WebPSSM subtype C sinisi (PSSM-C; <https://indra.mullins.microbiol.washington.edu/webpssm>) and PhenoSeq (www.burnet.edu.au/phenoseq) were used to predict the major coreceptor usage of participant - derived sequences.^{40–42} All three tools predict the major coreceptor usage of Env V3 sequences, however, they are not optimized to distinguish R5X4-tropic variants from X4-tropic variants. Here, we designated variants as R5 if they were predicted to use CCR5, and R5X4/X4 if they were predicted to use CXCR4 in combination with CCR5 or alone. Where the algorithms differed in predictions, we characterized an Env as R5 or R5X4/X4 based on two concordant algorithms.

Identification of N-linked glycosylation sites in the Env V3 loop sequences

Using the N-GlycoSite tool (<https://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html>), the location and number of putative N-linked glycosylation sequons (tripeptide N-X-S and N-X-T, where X is any amino acid except proline) in V3 loop sequences was recorded.⁶¹

Net charge calculation

Positively charged amino acids (lysine, arginine, and histidine) were scored as +1, and negatively charged amino acids (aspartate and glutamate) were scored as -1. The net charge of the Env V3 was the sum of all positive and negative charges within the V3 sequence.

Statistical analyses

Two-tailed, unpaired nonparametric t-tests were performed using GraphPad Prism version 5 for Windows (GraphPad Software, San Diego, CA) unless otherwise stated.

Results

Clinical features of individuals harboring HIV-1 and presenting with CM

The detailed demographic and clinical characteristics of HIV-1 infected individuals with CM in the parent study have been described in detail elsewhere.⁴⁷ Here, we randomly selected eight males and eight females infected with HIV-1 and presenting with CM, with matched plasma and CSF samples available for HIV-1 genotypic analyses. The samples were collected prior to initiation of cART. Demographic and clinical characteristics of these study participants are summarized in Table 1. The median age of the participants was 29.5 years (range: 23–45), with no significant difference between the median age of males and females ($p = 0.46$), and the median CD4⁺ T-cell count was 27.5 (interquartile range: 10–72 cells/ μ L). The median plasma and CSF HIV-1 viral load of the participants were 5.1 (interquartile range: 4.72–5.2) log₁₀ copies/mL and 4.73 (range: 4.3–4.7) log₁₀ copies/mL, respectively, and were not significantly different ($p = 0.13$). There was no significant difference between male and female plasma HIV-1 viral load ($p = 0.23$), and no significant difference between male and female CSF HIV-1 viral load ($p = 0.88$). Two participants, CM032 and CM117, did not have detectable white blood cells in the CSF, whereas all other participants exhibited

pleocytosis (>5 CSF white blood cells/ μL). We observed a positive correlation between the CD4^+ T-cell count and white blood cell count in the CSF (spearman $r = 0.7$, $p = 0.0026$).

The phylogenetic relationships of HIV-1 variants

We amplified and sequenced a total of 251 unique *env* genes from the plasma and CSF compartments of 16 participants, using either single-genome amplification (SGA) and sequencing, or bulk PCR, cloning and sequencing methods. The number of sequences generated per compartment of each participant using the two methods is summarized in Table 2. Phylogenetic analysis, performed as described in Materials and methods, revealed that all the sequences generated were subtype C. As expected, plasma and/or CSF sequences of each individual participant formed a unique cluster. All of the eleven participants with available matched plasma and CSF sequences had evidence of inter-compartment mixing of *env* sequences (Figure 1).

Genetic compartmentalization between anatomical compartments is more accurately determined using SGA-derived sequences than bulk PCR-derived sequences, as the former limits or eliminates taq-induced recombination, nucleotide misincorporation, template resampling and cloning bias.⁴⁹ We sought to establish whether there was genetic compartmentalization of HIV-1C sequences between the plasma and CSF or not of two participants (CM089 and CM112) where we generated at least 10 sequences from both of their compartments by SGA. For both participants, inter-compartment mixing of *env* sequences was evident (Figure 2). The Slatkin-Maddison test revealed seven and 10 inter-compartment migrations in participants CM089 and CM112, respectively. Additionally, neither CM089 nor CM112 demonstrated compartmentalization of *env* sequences (p -value 0.19, and 0.35 respectively). This result was confirmed with the Hudson nearest-neighbor test (p -value 0.34 and 0.58 for CM089 and CM112, respectively).

Predicted coreceptor usage of HIV-1C in plasma and CSF

To get better insight into coreceptor usage by HIV-1 subtype C variants in end-stage disease, we used three publicly available CPAs (Geno2pheno, PSSM-C and PhenoSeq) to generate a consensus prediction of the coreceptor usage of unique V3 sequences obtained in this study (Table 3).

We also assessed whether predicted R5 variants predominate over CXCR4-using variants in plasma and CSF samples (Figure 3). From the 14 plasma samples where we generated HIV-1 *env* genes, we identified that seven of 14 (50%) participants harbored R5 variants only, whereas one participant (7.1%) harbored only CXCR4-using variants, and six participants (42.9%) harbored a combination of R5 and CXCR4-using variants. Thus, based on the plasma samples only, R5 variants featured in 13 participants (92.9%) and CXCR4-using variants were detected in seven participants (50%). When we assessed the predicted coreceptor usage of variants in the 13 CSF samples where we generated sequences, we identified that nine of 13 (69.2%) CSF samples harbored R5 variants only, whereas two (15.4%) CSF samples each harbored R5 and CXCR4-using, and CXCR4-using variants only. Therefore, R5 variants featured in 11/13 (84.6%) CSF samples, while CXCR4-using variants were in 4/13 (30.8%) CSF samples. These results suggest that R5 variants

predominate in plasma as well as CSF samples, and that CXCR4 usage is more common in plasma than in the CSF of HIV-1C infected participants with CM. Overall, seven of the 16 (~43.8%) participants included in this analysis harbored CXCR4-using variants in plasma and/or CSF.

We next assessed whether matched plasma and CSF samples have HIV-1C variants with concordant coreceptor usage or not (Table 4). Seven of 11 (63.6%) participants with matched plasma and CSF sequences displayed concordant predicted coreceptor tropism of HIV-1C. A further two of 11 (18.2%) participants had discordance of predicted coreceptor tropism between plasma and CSF, where both R5 and CXCR4-using were detected in plasma but only CXCR4-using or R5 variants were detected in CSF. In two of 11 (18.2%) participants (CM019 and CM032), the concordance of coreceptor usage between the plasma and CSF was indeterminate as we only had one CSF sequence compared to multiple plasma sequences. Our results show that concordant predicted coreceptor tropism between matched plasma and CSF compartments predominates over discordance in individuals with CM.

Coreceptor usage determinants of HIV-1C Env

Understanding the determinants of CCR5 or CXCR4 usage by HIV-1C is of clinical relevance given that a coreceptor antagonist, maraviroc, is FDA approved and available to individuals who harbor R5 variants only.^{38, 62} Previously, it was shown that HIV-1C CXCR4-using variants have specific Env third variable loop (V3) properties which distinguish them from R5 variants. CXCR4-using HIV-1C variants have a V3 loop with greater amino acid variability, increased net charge, increased length, increased frequency of insertions, GPGQ crown motif substitutions, lower glycosylation site numbers, and charged residues at position 11 and/or 25 of the V3.^{15, 21, 63–66} Given that we identified a high prevalence of predicted CXCR4 usage by variants in this study, we sought to identify genetic properties of the major determinant of coreceptor usage, the V3 loop, which confer the ability to use CCR5 and/or CXCR4.

We determined the following additional properties of the V3 of all the sequences generated in our study: the overall net charge, number of potential N-linked glycosylation sites (PNGS), residues at position 11 and 25, and length (Table 3). Although 251 unique *env* genes were available in this study, only 58 unique V3 sequences were identified and aligned with a reference South African HIV-1C sequence (Figure 4). The three CPAs all agreed on the predicted coreceptor usage for 48 of 58 (82.8%) unique V3 sequences, but in 10 of 58 (17.2%) instances only two of the three CPAs agreed with each other.

As expected, the net charge of variants predicted to use CXCR4 (median 5; range 3-9) was significantly higher ($p < 0.0001$) than that of variants predicted to use CCR5 only (median 4; range 2-5) (Figure 5). CXCR4-using variants of participants CM019, CM050 and CM052 did not have a positively charged residue at position 11 and/or 25 of the V3, whereas participants CM032, CM070, CM098 and CM108 harbored at least one variant with a histidine, arginine or lysine at position 11 and/or 25. No variant predicted as R5 possessed a charged residue at position 11 and/or 25. This suggests that a charged residue at position 11 and/or 25 may be used to confer usage of CXCR4 for some CXCR4-using variants. Additionally, all CXCR4-using variants (except three from participant CM032 and one from

CM052) had a charge residue at position 16 and/or 18 of the V3, within the crown motif. R5 variants maintained a GPGQ crown motif generally.

We then identified multiple sequences from participant CM019 and one from participant CM032 which did not have PNGS and were predicted to use CXCR4, while every other sequence had one PNGS only at position 6 of the V3. This suggested that an absence of a PNGS in the V3 may contribute to the usage of CXCR4, but is not compulsory for CXCR4 usage. The range of V3 length for CXCR4-using and CCR5 variants was 28-37 and 34-35 amino acids, respectively, both with a median of 35 (Figure 5).

Rare mutation of the V3

The V3 loop of HIV-1 Env is a region important for function and maintains signature sequences according to subtype. The isoleucine at position 14 of the V3 (position 309 according to HXB2 numbering or I309) is important for the structure of the V3 loop, and is conserved in approximately 99% of HIV-1C sequences.⁶⁷ I309 also modulates the exposure of the CD4 binding site.⁶⁸ Notably, five unique V3 loop sequences of participant CM108 had a rare replacement of I309 with a phenylalanine residue, while every other sequence analyzed had an isoleucine at that position including the reference South African HIV-1C sequence (Ref.C.ZA) (Figure 4). The I309F residue was not specific to variants predicted to use CCR5, as it was present in variants predicted to use CXCR4 as well.

Discussion

Although others have studied HIV-1C *env* genes in individuals with advanced HIV disease and AIDS-defining illnesses^{15, 17, 64, 69}, genetic information of HIV-1C in matched blood and the CNS compartments of individuals with CM is sparse. In this study, we characterized the phylogenetic relatedness, and predicted coreceptor usage of plasma and CSF-derived HIV-1 variants in individuals with CM.

The sampling of HIV-1 from the CSF has been shown to be an adequate surrogate of sampling virus from the brain.^{29, 70} Equilibration, partial or complete, and compartmentalization of viral sequences between plasma and CSF have been described previously.^{24, 30, 33, 34} The occurrence of extreme genotypic compartmentalization of HIV-1 in the CNS, in particular, is suggestive of independent replication and evolution of viruses relative to the blood, and this is strongly associated with the development of neurocognitive impairment.^{31–33} The implications of equilibration or compartmentalization of HIV-1 sequences between blood and the CNS are unclear for individuals with co-infections. Here, we observed frequent intermixing of HIV-1C plasma and CSF-derived sequences in all 11 participants, such that plasma and CSF sequences did not form separate clusters but were phylogenetically interspersed among each other. Additionally, using the Slatkin-Maddison and Hudson nearest-neighbor compartmentalization algorithms on sequences from two participants CM089 and CM112, we observed an absence of compartmentalization of CSF-derived sequences in both participants. The majority of the participants in this study exhibited pleocytosis, a marker of CNS inflammation which is associated with HIV-1 trafficking between peripheral blood and the CNS⁷¹, therefore it is probable that *Cryptococcus neoformans* and/or HIV-1C co-infection led to a breach of the

blood-brain barrier which allowed movement of viruses between the peripheral blood and CNS compartments. Participants CM032 and CM117 did not display pleocytosis, however, though they had inter-compartment mixing of sequences suggesting that pleocytosis is not a prerequisite for genetic intermixing between blood and the CNS. Because we did not evaluate longitudinal samples, we also could not rule out a previous pleocytosis event that may have led to an absence of viral genetic segregation in these participants. Nevertheless, these results supported our hypothesis where we expected that inter-compartment mixing would arise in some individuals with CM. It is possible that immunological compartmentalization was a recent event associated with the development of active cryptococcosis, and therefore did not allow enough time for detectable independent viral evolution.

R5 variants establish infection in the CNS compartment^{23, 24}, but R5, R5X4 and X4 have been isolated from the brain later in disease.⁷² We observed that concordance of HIV-1 coreceptor usage predominated in matched plasma and CSF samples. Our results were in agreement with other studies where phenotypic assays^{35, 36} and CPAs³⁷ were employed, showing that concordance of HIV-1 coreceptor usage predominated but discordance of HIV-1 between the plasma and CSF occurred in 10.9% to 35.7% of matched samples. Additionally, we showed that plasma may harbor CXCR4-using variants while the CSF harbors R5 variants only as described previously.^{36, 37} These results suggest that discordance of coreceptor usage between compartments was reflective of the functional compartmentalization of HIV-1, as others have shown that *env* sequences in the CNS may be compartmentalized.⁷³

Previously, it was shown that CXCR4-using HIV-1C has greater amino acid variability, increased net charge, increased length, increased frequency of insertions, GPGQ crown motif substitutions, lower glycosylation site numbers, and charged residues at position 11 and/or 25 of the V3 relative to R5 variants.^{21, 63–66} In this study, we investigated whether the HIV-1C R5 and CXCR4-using variants predicted by CPAs possessed the same determinants of coreceptor usage as variants from previous studies. We identified that many previously described V3 properties specific to HIV-1C R5 and CXCR4-using variants are consistent with variants in our study. In summary, the predicted CXCR4-using variants in our study had, compared to R5 variants, V3 sequences with higher net charge, two amino acids insertions upstream of the crown motif, an absence of PNGS, a charged residue at position 11/and 25, and alterations of the GPGQ crown motif.

Changes in position 14 of the V3 (position 309 according to HXB2 numbering) affect the structure of the loop⁶⁷, and the exposure of the CD4 binding site.⁶⁸ We identified that sequences from participant CM108 possessed a change of the conserved isoleucine at V3 position 14 (I309) to phenylalanine (F309). Furthermore, the F309 residue was present in variants predicted as R5 and CXCR4-using, suggesting that this substitution was not unique to any particular coreceptor usage phenotype. Given that I309 was shown to influence the structure of the V3 and modulate the exposure of the CD4 binding site, it suggests that variants in participant CM108 may have a unique entry and immune evasion profile compared to variants with the I309 residue harbored by other individuals with CM.

Reliable estimates of the prevalence of CXCR4 by HIV-1 are required to understand the evolution of the HIV-1C epidemic and to inform effective therapeutic interventions. HIV-1B has been shown to expand its usage of coreceptors in chronic and the end stages of disease. 10–13 On the other hand, HIV-1C has been observed to switch major coreceptor usage less frequently. 14, 15, 19–21 In some cases, a complete absence of CXCR4 usage by HIV-1C was observed in individuals with advanced disease, suggesting that the ability to switch major coreceptors for this subtype is rare or not essential. 17, 18 We showed that R5 variants predominate in both plasma as well as CSF samples, but that at least one CXCR4-using variant was present in plasma and/or CSF in ~43.8% of study-participants. Although we expected to detect CXCR4 usage by HIV-1C in some individuals in this study, considering that all had an AIDS defining condition, we observed a much higher frequency of CXCR4 usage than expected based on previous studies. It may be that the CPAs overestimated the predicted CXCR4 usage as described previously. 74–77 Coreceptor usage tests *in vitro* may be employed in future to corroborate our findings. It remains unclear, however, whether CXCR4-using variants are enriched in individuals depending on the co-infection they have, or that HIV-1C has evolved to use CXCR4 more over time. We also showed that 4/13 (30.8%) CSF samples harbored predicted CXCR4-using HIV-1C variants, indicating that different phenotypes besides R5 variants may circulate in the CNS in chronic to end-stage disease. The clinical importance of our findings may require further longitudinal evaluation considering that genetic intermixing between peripheral blood and the CNS may afford HIV-1 the ability to target cells of the CNS using CXCR4.

Our study was limited to a small cohort of 16 participants with highly similar clinical characteristics. We did not have a control group of participants infected with HIV-1 but not presenting with CM. We generated a low number of sequence for one or both compartments of some participants, therefore we may have missed minority lineages that are specific to the peripheral blood or CNS compartment. Another limitation of the study which may have impacted the results of our study was the use of bulk PCR amplification and cloning together with SGA to generate sequences. SGA has been shown to provide a better measure of viral variants, especially minority ones, in a quasispecies compared to bulk PCR and cloning, as described previously. 49 On the other hand, a different study showed that bulk PCR and cloning may capture viral diversity as SGA does. 78

Future longitudinal studies assessing differences of HIV-1 in blood and the CNS of individuals with and without CM will improve our understanding of the pathogenesis of HIV-1. Our results suggest that trafficking of HIV-1 variants between blood and the CNS is common, though the mode and the effect on HIV-1 neurotropism is not completely clear. We also established using CPAs that maraviroc use would not be suitable for approximately 44% of individuals with CM, suggesting an alternative mode of therapy must be optimized for them if standard regimens fail.

Acknowledgements

This research was funded by the Howard Hughes Medical Institute, the South African Department of Science and Technology/National Research Foundation Research Chairs Initiative and the Victor Daitz Foundation. Partial funding was provided by the Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE), a DELTAS Africa Initiative [grant # DEL-15-006]. The DELTAS Africa Initiative is an independent funding scheme

of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [grant # 107752/Z/15/Z] and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government. We extend our gratitude to Professor Paul Gorry, Head of the HIV Molecular Pathogenesis Laboratory (Burnet Institute), for providing us with the pSVIII-Env plasmid. Finally, we thank and acknowledge the study participants.

References

1. Alkhatib G, Combadiere C, Broder CC, et al. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science*. 1996 Jun 28; 272(5270):1955–1958. [PubMed: 8658171]
2. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*. 1996 May 10; 272(5263):872–877. [PubMed: 8629022]
3. Deng H, Liu R, Ellmeier W, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*. 1996 Jun 20; 381(6584):661–666. [PubMed: 8649511]
4. Xiao L, Owen SM, Goldman I, et al. CCR5 coreceptor usage of non-syncytium-inducing primary HIV-1 is independent of phylogenetically distinct global HIV-1 isolates: delineation of consensus motif in the V3 domain that predicts CCR-5 usage. *Virology*. 1998 Jan 5; 240(1):83–92. [PubMed: 9448692]
5. Verrier F, Borman AM, Brand D, Girard M. Role of the HIV type 1 glycoprotein 120 V3 loop in determining coreceptor usage. *AIDS Res Hum Retroviruses*. 1999 May 20; 15(8):731–743. [PubMed: 10357469]
6. Cashin K, Sterjovski J, Harvey KL, Ramsland PA, Churchill MJ, Gorro PR. Covariance of charged amino acids at positions 322 and 440 of HIV-1 Env contributes to coreceptor specificity of subtype B viruses, and can be used to improve the performance of V3 sequence-based coreceptor usage prediction algorithms. *PLoS One*. 2014; 9(10):e109771. [PubMed: 25313689]
7. Cho MW, Lee MK, Carney MC, Berson JF, Doms RW, Martin MA. Identification of determinants on a dualtropic human immunodeficiency virus type 1 envelope glycoprotein that confer usage of CXCR4. *J Virol*. 1998 Mar; 72(3):2509–2515. [PubMed: 9499115]
8. Smyth RJ, Yi Y, Singh A, Collman RG. Determinants of entry cofactor utilization and tropism in a dualtropic human immunodeficiency virus type 1 primary isolate. *J Virol*. 1998 May; 72(5):4478–4484. [PubMed: 9557745]
9. Huang W, Toma J, Fransen S, et al. Coreceptor tropism can be influenced by amino acid substitutions in the gp41 transmembrane subunit of human immunodeficiency virus type 1 envelope protein. *J Virol*. 2008 Jun; 82(11):5584–5593. [PubMed: 18353956]
10. Scarlatti G, Tresoldi E, Bjorndal A, et al. In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. *Nat Med*. 1997 Nov; 3(11):1259–1265. [PubMed: 9359702]
11. Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in coreceptor use correlates with disease progression in HIV-1-infected individuals. *J Exp Med*. 1997 Feb 17; 185(4):621–628. [PubMed: 9034141]
12. Koot M, Vos AH, Keet RP, et al. HIV-1 biological phenotype in long-term infected individuals evaluated with an MT-2 cocultivation assay. *AIDS*. 1992 Jan; 6(1):49–54. [PubMed: 1543566]
13. Koot M, Keet IP, Vos AH, et al. Prognostic value of HIV-1 syncytium-inducing phenotype for rate of CD4+ cell depletion and progression to AIDS. *Ann Intern Med*. 1993 May 1; 118(9):681–688. [PubMed: 8096374]
14. Abebe A, Demissie D, Goudsmit J, et al. HIV-1 subtype C syncytium- and non-syncytium-inducing phenotypes and coreceptor usage among Ethiopian patients with AIDS. *AIDS*. 1999 Jul 30; 13(11):1305–1311. [PubMed: 10449282]
15. Jakobsen MR, Cashin K, Roche M, et al. Longitudinal Analysis of CCR5 and CXCR4 Usage in a Cohort of Antiretroviral Therapy-Naive Subjects with Progressive HIV-1 Subtype C Infection. *PLoS One*. 2013; 8(6)

16. Choge I, Cilliers T, Walker P, et al. Genotypic and phenotypic characterization of viral isolates from HIV-1 subtype C-infected children with slow and rapid disease progression. *AIDS Res Hum Retroviruses*. 2006 May; 22(5):458–465. [PubMed: 16706624]
17. Morris L, Cilliers T, Bredell H, Phoswa M, Martin DJ. CCR5 is the major coreceptor used by HIV-1 subtype C isolates from patients with active tuberculosis. *AIDS Res Hum Retroviruses*. 2001 May 20; 17(8):697–701. [PubMed: 11429110]
18. Cecilia D, Kulkarni SS, Tripathy SP, Gangakhedkar RR, Paranjape RS, Gadkari DA. Absence of coreceptor switch with disease progression in human immunodeficiency virus infections in India. *Virology*. 2000 Jun 5; 271(2):253–258. [PubMed: 10860879]
19. Connell BJ, Michler K, Capovilla A, Venter WD, Stevens WS, Papathanasopoulos MA. Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa. *AIDS*. 2008 Apr 23; 22(7):896–899. [PubMed: 18427209]
20. Lin NH, Smeaton LM, Giguel F, et al. Prevalence and clinical associations of CXCR4-using HIV-1 among treatment-naïve subtype C-infected women in Botswana. *J Acquir Immune Defic Syndr*. 2011 May 1; 57(1):46–50. [PubMed: 21346588]
21. Cilliers T, Nhlapo J, Coetzer M, et al. The CCR5 and CXCR4 coreceptors are both used by human immunodeficiency virus type 1 primary isolates from subtype C. *J Virol*. 2003 Apr; 77(7):4449–4456. [PubMed: 12634405]
22. Sonnerborg A, Nybom R, Britton S, et al. Detection of cell-free human immunodeficiency virus in cerebrospinal fluid by using immune scanning electron microscopy. *J Infect Dis*. 1989 Jun; 159(6):1037–1041. [PubMed: 2498434]
23. Schnell G, Price RW, Swanstrom R, Spudich S. Compartmentalization and clonal amplification of HIV-1 variants in the cerebrospinal fluid during primary infection. *J Virol*. 2010 Mar; 84(5):2395–2407. [PubMed: 20015984]
24. Sturdevant CB, Joseph SB, Schnell G, Price RW, Swanstrom R, Spudich S. Compartmentalized replication of R5 T cell-tropic HIV-1 in the central nervous system early in the course of infection. *PLoS Pathog*. 2015 Mar; 11(3):e1004720. [PubMed: 25811757]
25. Brew BJ, Rosenblum M, Cronin K, Price RW. AIDS dementia complex and HIV-1 brain infection: clinical-virological correlations. *Ann Neurol*. 1995 Oct; 38(4):563–570. [PubMed: 7574452]
26. Ellis RJ, Gamst AC, Capparelli E, et al. Cerebrospinal fluid HIV RNA originates from both local CNS and systemic sources. *Neurology*. 2000 Feb 22; 54(4):927–936. [PubMed: 10690988]
27. Di Stefano M, Monno L, Fiore JR, et al. Neurological disorders during HIV-1 infection correlate with viral load in cerebrospinal fluid but not with virus phenotype. *AIDS*. 1998 May 7; 12(7):737–743. [PubMed: 9619805]
28. Price RW, Staprans S. Measuring the "viral load" in cerebrospinal fluid in human immunodeficiency virus infection: window into brain infection? *Ann Neurol*. 1997 Nov; 42(5):675–678. [PubMed: 9392565]
29. Strain MC, Letendre S, Pillai SK, et al. Genetic composition of human immunodeficiency virus type 1 in cerebrospinal fluid and blood without treatment and during failing antiretroviral therapy. *J Virol*. 2005 Feb; 79(3):1772–1788. [PubMed: 15650202]
30. Pillai SK, Pond SL, Liu Y, et al. Genetic attributes of cerebrospinal fluid-derived HIV-1 env. *Brain*. 2006 Jul; 129(Pt 7):1872–1883. [PubMed: 16735456]
31. Ritola K, Robertson K, Fiscus SA, Hall C, Swanstrom R. Increased human immunodeficiency virus type 1 (HIV-1) env compartmentalization in the presence of HIV-1-associated dementia. *J Virol*. 2005 Aug; 79(16):10830–10834. [PubMed: 16051875]
32. Harrington PR, Schnell G, Letendre SL, et al. Cross-sectional characterization of HIV-1 env compartmentalization in cerebrospinal fluid over the full disease course. *AIDS*. 2009 May 15; 23(8):907–915. [PubMed: 19414991]
33. Schnell G, Joseph S, Spudich S, Price RW, Swanstrom R. HIV-1 replication in the central nervous system occurs in two distinct cell types. *PLoS Pathog*. 2011 Oct; 7(10):e1002286. [PubMed: 22007152]
34. Sturdevant CB, Dow A, Jabara CB, et al. Central nervous system compartmentalization of HIV-1 subtype C variants early and late in infection in young children. *PLoS Pathog*. 2012 Dec; 8(12):e1003094. [PubMed: 23300446]

35. Spudich SS, Huang W, Nilsson AC, et al. HIV-1 chemokine coreceptor utilization in paired cerebrospinal fluid and plasma samples: a survey of subjects with viremia. *J Infect Dis.* 2005 Mar 15; 191(6):890–898. [PubMed: 15717264]
36. Karlsson U, Antonsson L, Repits J, et al. Mode of coreceptor use by R5 HIV type 1 correlates with disease stage: a study of paired plasma and cerebrospinal fluid isolates. *AIDS Res Hum Retroviruses.* 2009 Dec; 25(12):1297–1305. [PubMed: 20001314]
37. Soulie C, Tubiana R, Simon A, et al. Presence of HIV-1 R5 viruses in cerebrospinal fluid even in patients harboring R5X4/X4 viruses in plasma. *J Acquir Immune Defic Syndr.* 2009 May 1; 51(1): 60–64. [PubMed: 19390328]
38. Fatkenheuer G, Pozniak AL, Johnson MA, et al. Efficacy of short-term monotherapy with maraviroc, a new CCR5 antagonist, in patients infected with HIV-1. *Nat Med.* 2005 Nov; 11(11): 1170–1172. [PubMed: 16205738]
39. Aiamkitsumrit B, Dampier W, Antell G, et al. Bioinformatic analysis of HIV-1 entry and pathogenesis. *Curr HIV Res.* 2014; 12(2):132–161. [PubMed: 24862329]
40. Jensen MA, Li FS, van 't Wout AB, et al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol.* 2003 Dec; 77(24):13376–13388. [PubMed: 14645592]
41. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. *Nat Biotechnol.* 2007 Dec; 25(12):1407–1410. [PubMed: 18066037]
42. Cashin K, Gray LR, Harvey KL, et al. Reliable genotypic tropism tests for the major HIV-1 subtypes. *Sci Rep.* 2015; 5:8543. [PubMed: 25712827]
43. Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, Chiller TM. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. *AIDS.* 2009 Feb 20; 23(4):525–530. [PubMed: 19182676]
44. Jarvis JN, Bicanic T, Loyse A, et al. Determinants of mortality in a combined cohort of 501 patients with HIV-associated Cryptococcal meningitis: implications for improving outcomes. *Clin Infect Dis.* 2014 Mar; 58(5):736–745. [PubMed: 24319084]
45. Mwaba P, Mwansa J, Chintu C, et al. Clinical presentation, natural history, and cumulative death rates of 230 adults with primary cryptococcal meningitis in Zambian AIDS patients treated under local conditions. *Postgrad Med J.* 2001 Dec; 77(914):769–773. [PubMed: 11723315]
46. Naranbhai V, Chang CC, Durgiah R, et al. Compartmentalization of innate immune responses in the central nervous system during cryptococcal meningitis/HIV coinfection. *AIDS.* 2014 Mar 13; 28(5):657–666. [PubMed: 24451162]
47. Chang CC, Dorasamy AA, Gosnell BI, et al. Clinical and mycological predictors of cryptococcosis-associated immune reconstitution inflammatory syndrome. *AIDS.* 2013 Aug 24; 27(13):2089–2099. [PubMed: 23525034]
48. Chang, CC, Kangethe, R, Omarjee, S. , et al. Open Forum Infect Dis. Vol. 4. Spring; 2017. Relationship of Human Immunodeficiency Virus Viral Load in Cerebrospinal Fluid and Plasma in Patients Co-infected With Cryptococcal Meningitis. ofx032
49. Salazar-Gonzalez JF, Bailes E, Pham KT, et al. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J Virol.* 2008 Apr; 82(8):3952–3970. [PubMed: 18256145]
50. Palmer S, Kearney M, Maldarelli F, et al. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. *J Clin Microbiol.* 2005 Jan; 43(1):406–413. [PubMed: 15635002]
51. Archary D, Gordon ML, Green TN, Coovadia HM, Goulder PJ, Ndung'u T. HIV-1 subtype C envelope characteristics associated with divergent rates of chronic disease progression. *Retrovirology.* 2010; 7:92. [PubMed: 21050445]
52. Gao F, Morrison SG, Robertson DL, et al. Molecular cloning and analysis of functional envelope genes from human immunodeficiency virus type 1 sequence subtypes A through G. The WHO and NIAID Networks for HIV Isolation and Characterization. *J Virol.* 1996 Mar; 70(3):1651–1667. [PubMed: 8627686]

53. Roche M, Jakobsen MR, Sterjovski J, et al. HIV-1 escape from the CCR5 antagonist maraviroc associated with an altered and less-efficient mechanism of gp120-CCR5 engagement that attenuates macrophage tropism. *J Virol*. 2011 May; 85(9):4330–4342. [PubMed: 21345957]
54. Ohagen A, Devitt A, Kunstman KJ, et al. Genetic and functional analysis of full-length human immunodeficiency virus type 1 env genes derived from brain and blood of patients with AIDS. *J Virol*. 2003 Nov; 77(22):12336–12345. [PubMed: 14581570]
55. Sterjovski J, Churchill MJ, Ellett A, et al. Asn 362 in gp120 contributes to enhanced fusogenicity by CCR5-restricted HIV-1 envelope glycoprotein variants from patients with AIDS. *Retrovirology*. 2007 Dec 12; 4:89. [PubMed: 18076768]
56. Rousseau CM, Birditt BA, McKay AR, et al. Large-scale amplification, cloning and sequencing of near full-length HIV-1 subtype C genomes. *J Virol Methods*. 2006 Sep; 136(1–2):118–125. [PubMed: 16701907]
57. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016 Jul; 33(7):1870–1874. [PubMed: 27004904]
58. Kumar SB, Handelsman SK, Voronkin I, et al. Different regions of HIV-1 subtype C env are associated with placental localization and in utero mother-to-child transmission. *J Virol*. 2011 Jul; 85(14):7142–7152. [PubMed: 21543508]
59. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003 Oct; 52(5):696–704. [PubMed: 14530136]
60. Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*. 2005 Mar 1; 21(5):676–679. [PubMed: 15509596]
61. Zhang M, Gaschen B, Blay W, et al. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology*. 2004 Dec; 14(12):1229–1246. [PubMed: 15175256]
62. Gulick RM, Lalezari J, Goodrich J, et al. Maraviroc for previously treated patients with R5 HIV-1 infection. *N Engl J Med*. 2008 Oct 2; 359(14):1429–1441. [PubMed: 18832244]
63. Cashin K, Gray LR, Jakobsen MR, Sterjovski J, Churchill MJ, Gorry PR. CoRSeqV3-C: a novel HIV-1 subtype C specific V3 sequence based coreceptor usage prediction algorithm. *Retrovirology*. 2013; 10:24. [PubMed: 23446039]
64. Coetzer M, Cilliers T, Ping LH, Swanstrom R, Morris L. Genetic characteristics of the V3 region associated with CXCR4 usage in HIV-1 subtype C isolates. *Virology*. 2006 Dec 5–20; 356(1–2): 95–105. [PubMed: 16942785]
65. Zhang H, Tully DC, Zhang T, Moriyama H, Thompson J, Wood C. Molecular determinants of HIV-1 subtype C coreceptor transition from R5 to R5X4. *Virology*. 2010 Nov 10; 407(1):68–79. [PubMed: 20797755]
66. Lin NH, Becerril C, Giguel F, et al. Env sequence determinants in CXCR4-using human immunodeficiency virus type-1 subtype C. *Virology*. 2012 Nov 25; 433(2):296–307. [PubMed: 22954962]
67. Stanfield RL, Gorny MK, Zolla-Pazner S, Wilson IA. Crystal structures of human immunodeficiency virus type 1 (HIV-1) neutralizing antibody 2219 in complex with three different V3 peptides reveal a new binding mode for HIV-1 cross-reactivity. *J Virol*. 2006 Jun; 80(12):6093–6105. [PubMed: 16731948]
68. Lynch RM, Rong R, Li B, et al. Subtype-specific conservation of isoleucine 309 in the envelope V3 domain is linked to immune evasion in subtype C HIV-1 infection. *Virology*. 2010 Aug 15; 404(1):59–70. [PubMed: 20494390]
69. Michler K, Connell BJ, Venter WD, Stevens WS, Capovilla A, Papathanasopoulos MA. Genotypic characterization and comparison of full-length envelope glycoproteins from South African HIV type 1 subtype C primary isolates that utilize CCR5 and/or CXCR4. *AIDS Res Hum Retroviruses*. 2008 May; 24(5):743–751. [PubMed: 18507530]
70. McArthur JC, McClernon DR, Cronin MF, et al. Relationship between human immunodeficiency virus-associated dementia and viral load in cerebrospinal fluid and brain. *Ann Neurol*. 1997 Nov; 42(5):689–698. [PubMed: 9392567]

71. Smith DM, Zarate S, Shao H, et al. Pleocytosis is associated with disruption of HIV compartmentalization between blood and cerebral spinal fluid viral populations. *Virology*. 2009 Mar 1; 385(1):204–208. [PubMed: 19100592]
72. Gorry PR, Bristol G, Zack JA, et al. Macrophage tropism of human immunodeficiency virus type 1 isolates from brain and lymphoid tissues predicts neurotropism independent of coreceptor specificity. *J Virol*. 2001 Nov; 75(21):10073–10089. [PubMed: 11581376]
73. Korber BT, Kunstman KJ, Patterson BK, et al. Genetic differences between blood- and brain-derived viral sequences from human immunodeficiency virus type 1-infected patients: evidence of conserved elements in the V3 region of the envelope protein of brain-derived sequences. *J Virol*. 1994 Nov; 68(11):7467–7481. [PubMed: 7933130]
74. Ceresola ER, Nozza S, Sampaolo M, et al. Performance of commonly used genotypic assays and comparison with phenotypic assays of HIV-1 coreceptor tropism in acutely HIV-1-infected patients. *J Antimicrob Chemother*. 2015 May; 70(5):1391–1395. [PubMed: 25608585]
75. Mulinge M, Lemaire M, Servais JY, et al. HIV-1 tropism determination using a phenotypic Env recombinant viral assay highlights overestimation of CXCR4-usage by genotypic prediction algorithms for CRF01_AE and CRF02_AG [corrected]. *PLoS One*. 2013; 8(5):e60566. [PubMed: 23667426]
76. Raymond S, Delobel P, Mavigner M, et al. CXCR4-using viruses in plasma and peripheral blood mononuclear cells during primary HIV-1 infection and impact on disease progression. *AIDS*. 2010 Sep 24; 24(15):2305–2312. [PubMed: 20808203]
77. Seager I, Travers SA, Leeson MD, et al. Coreceptor usage, diversity, and divergence in drug-naïve and drug-exposed individuals from Malawi, infected with HIV-1 subtype C for more than 20 years. *AIDS Res Hum Retroviruses*. 2014 Oct; 30(10):975–983. [PubMed: 24925099]
78. Jordan MR, Kearney M, Palmer S, et al. Comparison of standard PCR/cloning to single genome sequencing for analysis of HIV-1 populations. *J Virol Methods*. 2010 Sep; 168(1–2):114–120. [PubMed: 20451557]

Figure 1. A neighbor-joining phylogenetic tree of participant-derived HIV-1 envelope gene sequences and a panel of reference sequences from different subtypes.

All of the sequences from individual study participants share the same color branches, and reference subtype sequences have grey branches. Closed circles and triangles at the tip of each branch indicate predicted R5 and CXCR4-using plasma-derived sequences, respectively. Open circles and triangles indicate predicted R5 and CXCR4-using cerebrospinal fluid-derived sequences, respectively. No inter-participant mixing of sequences was observed, suggesting an absence of inter-participant sequence contamination. Furthermore, all of the participant-derived sequences clustered with reference subtype C

sequences while all other reference sequence subtypes clustered together but separately from subtype C sequences.

Figure 2. A maximum-likelihood phylogenetic tree of HIV-1 subtype C envelope variants in study participant CM089 (A) and CM112 (B).

Envelope gene sequences labeled with red and blue circles were derived from plasma and cerebrospinal fluid, respectively. Bootstrap values greater than 70% are shown at their respective nodes and the genetic distance scale bar reflects the number of nucleotide substitutions per site.

Figure 3. Predicted coreceptor usage in plasma and CSF samples.

The number of plasma and CSF samples harboring R5, R5 and CXCR4-using or CXCR4-using variants are shown as stacked columns.

Figure 4. An amino acid sequence alignment of unique HIV-1C envelope V3 sequences from plasma and cerebrospinal fluid-derived viruses of antiretroviral therapy naïve individuals living with cryptococcal meningitis (CM).

A reference South African HIV-1C envelope (Ref.C.ZA) V3 sequence was aligned with CM participant-derived sequences using MUSCLE.⁷⁹ A star beside the sequence name indicates a variant that circulates in CSF, whereas the absence of a star beside the sequence name indicates a variant that circulates in plasma. The first and second number in the round brackets indicates the number of identical V3 sequences in other variants circulating in plasma and CSF, respectively. The predicted coreceptor usage of each V3 sequence is

indicated: “R5” indicates it was predicted to use CCR5 only, and “R5X4/X4” indicates it was predicted to use CXCR4 in combination with CCR5 or alone. Amino acids identical to the reference sequence are shown as dots, disagreements are indicated as bold lettering, dashes indicate deletions, and the number to the right of the amino acid sequence indicates the length of the V3. The light blue bar shows position 6 of the V3 where potential N-linked glycosylation sites are located, the orange bar shows amino acids at position 11, the pink bar shows where two-amino acid insertions occur, the dark blue bar shows the crown motif sequences and the green bar shows amino acids at position 25 of the V3.

Figure 5. The net charge and length of HIV-1 subtype C envelope V3.

The net charge (A) and length (B) of the V3 sequence of unique R5 (closed circles) and CXCR4-using (open circles) HIV-1C variants was compared. The red line indicates the median for each group.

Table 1
Demographic and clinical characteristics of the study participants.

ID ^a	Age	Sex	CD4 count ^b	Plasma VL ^c	CSF VL ^d	CSF WBC ^e
CM019	45	M	7	5.15	4.33	28
CM021	30	F	134	4.34	4.34	40
CM029	33	F	121	5.22	5.99	424
CM032	34	F	5	4.94	2.52	0
CM041	26	F	58	4.51	4.93	102
CM050	23	F	16	4.92	4.80	6
CM052	24	F	172	4.62	4.83	36
CM054	23	M	7	5.20	4.20	14
CM070	25	M	11	5.04	4.00	18
CM089	27	M	114	5.67	5.85	154
CM094	29	F	35	5.15	4.03	84
CM098	40	M	14	4.76	5.27	66
CM108	27	M	20	5.16	5.02	28
CM112	34	F	39	5.52	5.62	6
CM117	38	M	1	5.72	4.66	0
CM132	34	M	53	4.58	4.43	38

^aThe identification number of the study participant.

^bThe CD4⁺ T-cell count was recorded as cells/ μ L.

^{c,d}The cell-free HIV-1 RNA viral load (VL) in plasma and CSF was recorded as log₁₀ copies/mL.

^eThe CSF white blood cell (WBC) count was recorded as cells/ μ L.

Table 2
The number of sequences generated from the plasma and CSF of the study participants.

ID ^a	Plasma ^b		CSF ^c		Total
	Cloning	SGA	Cloning	SGA	
CM019	11	17	1	0	29
CM021	1	2	2	13	18
CM029	0	0	5	6	11
CM032	3	-	1	-	4
CM041	-	4	-	0	4
CM050	2	1	1	9	13
CM052	3	6	3	11	23
CM054	3	-	0	-	3
CM070	5	-	0	-	5
CM089	-	10	-	19	29
CM094	0	-	3	-	3
CM098	11	-	16	-	27
CM108	9	-	8	-	17
CM112	-	16	-	16	32
CM117	12	-	9	-	21
CM132	-	8	1	3	12
Total					251

^aThe identification number of the study participant.

^{b,c}Sequences were generated by the single-genome amplification (SGA) and sequencing, or bulk PCR, **cloning** and sequencing method. Dashes indicate where we did not attempt to generate sequences for this sample using this method, and zeros indicate where we were unable to amplify cDNA before the sample was completely consumed or we could not amplify an *env* gene.

Table 3

V3 loop properties of HIV-1C envelopes.

ID ^a	Compart.	Sequence ^b	G2P	PhenoSeq	PSSM-C	Consensus	Net charge	Pos. 11, 25	PNGS	Length
CM019	CSF	CM019C.B (4:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	5	SD	0	37
		CM019P2 (7:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	5	SD	0	37
	Plasma	CM019P.B.3 (7:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	5	SD	0	37
		CM019P.E (5:0)	R5	R5X4/X4	R5	R5	4	SD	1	35
		CM019P.L2 (0:0)	R5	R5	R5	R5	4	SD	1	35
CM021	CSF	CM019P_SGA14 (0:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	6	SD	0	37
		CM021C.7.3 (2:14)	R5	R5	R5	R5	3	SD	1	34
	Plasma	CM021P_SGA22 (0:0)	R5	R5	R5	R5	4	SD	1	35
		CM029C.B.3 (0:5)	R5	R5	R5	R5	4	SD	1	35
		CM029C.D.2 (0:1)	R5	R5	R5	R5	4	SD	1	35
CM029	CSF	CM029C.E.2 (0:1)	R5	R5	R5	R5	5	SD	1	35
		CM029C_SGA01 (0:0)	R5	R5	R5	R5	4	SD	1	35
		CM032C.6.4 (0:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	5	SR	1	35
	Plasma	CM032P.10 (1:0)	R5	R5X4/X4	R5X4/X4	R5X4/X4	4	SR	1	35
		CM032P.4 (0:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	9	HR	0	35
CM041	Plasma	CM041P_SGA01 (3:0)	R5	R5	R5	R5	4	SD	1	35
CM050	CSF	CM050C.5 (2:1)	R5	R5X4/X4	R5	R5	5	GD	1	35
		CM050C_SGA01 (1:7)	R5	R5X4/X4	R5X4/X4	R5X4/X4	6	S-	1	34
		CM052C.1 (5:7)	R5	R5	R5	R5	3	SD	1	35
	Plasma	CM052C_SGA04B (0:0)	R5	R5	R5	R5	4	SD	1	35
		CM052C_SGA10B (2:4)	R5	R5	R5	R5	4	SD	1	35
CM052	CSF	CM052P_SGA03 (0:0)	R5X4/X4	R5	R5X4/X4	R5X4/X4	4	SD	1	35
		CM052P_SGA06 (0:0)	R5	R5	R5	R5	5	SD	1	35
		CM054P.10.5 (2:0)	R5	R5	R5	R5	4	SD	1	35
	Plasma	CM070P.1 (1:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	7	SK	1	37
		CM070P.F.4 (1:0)	R5	R5	R5	R5	5	SD	1	35

ID ^a	Compart.	Sequence ^b	G2P	PhenoSeq	PSSM-C	Consensus	Net charge	Pos. 11, 25	PNGS	Length
CM089	CSF	CM070P.1.1 (0:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	8	NK	1	37
		CM089C_SGA01 (5:13)	R5	R5	R5	R5	4	SD	1	35
		CM089C_SGA07 (0:2)	R5	R5	R5X4/X4	R5	5	SG	1	35
		CM089C_SGA16 (0:0)	R5	R5	R5	R5	5	SG	1	35
		CM089C_SGA21 (3:0)	R5	R5	R5	R5	4	SD	1	35
	Plasma	CM089P_SGA05 (0:0)	R5	R5	R5	R5	5	SD	1	35
		CM089P_SGA09 (0:0)	R5	R5	R5	R5	5	SD	1	35
		CM094C.11.2 (0:2)	R5	R5	R5	R5	4	SD	1	34
		CM098C.1 (2:6)	R5	R5X4/X4	R5	R5	3	SE	1	35
		CM098C.C.2 (0:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	3	SE	1	28
CM098	CSF	CM098C.F.1 (5:0)	R5	R5	R5	R5	5	SA	1	35
		CM098C.F.2 (0:1)	R5	R5X4/X4	R5X4/X4	R5X4/X4	3	NE	1	35
		CM098C.H.1 (2:3)	R5	R5	R5	R5	4	SA	1	35
	Plasma	CM098C.J.4 (0:0)	R5	R5	R5	R5	4	SA	1	35
		CM098P.A.1 (1:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	6	SK	1	37
		CM108C.7.12 (0:6)	R5	R5	R5	R5	4	SD	1	35
CM108	CSF	CM108C.8.9 (0:0)	R5	R5	R5	R5	3	SD	1	35
		CM108P.1 (1:0)	R5X4/X4	R5X4/X4	R5X4/X4	R5X4/X4	7	RD	1	35
	Plasma	CM108PD.5 (4:0)	R5	R5X4/X4	R5X4/X4	R5X4/X4	5	SD	1	35
		CM108PK.1 (1:0)	R5	R5	R5X4/X4	R5	3	SD	1	35
	CSF	CM112C_SGA01 (11:11)	R5	R5	R5	R5	4	SD	1	35
		CM112C_SGA11 (1:0)	R5	R5	R5	R5	4	SD	1	35
CM112	CSF	CM112C_SGA12 (0:1)	R5	R5	R5	R5	4	SD	1	35
		CM112C_SGA13 (0:0)	R5	R5	R5	R5	3	SD	1	35
	Plasma	CM112P_SGA02 (1:0)	R5	R5	R5	R5	3	SD	1	35
		CM112P_SGA10 (0:0)	R5	R5	R5	R5	4	SE	1	35
	CSF	CM117C.A (12:7)	R5	R5	R5	R5	2	SD	1	35
		CM117C.H.1 (0:0)	R5	R5	R5	R5	2	SD	1	35

ID ^a	Compart.	Sequence ^b	G2P	PhenoSeq	PSSM-C	Consensus	Net charge	Pos. 11, 25	PNGS	Length
CMI32	CSF	CM132C.1 (3:2)	R5	R5	R5	R5	3	SD	1	35
		CM132C_SGA01 (0:0)	R5	R5	R5	R5	4	SD	1	35
	Plasma	CM132P_SGA02 (4:0)	R5	R5	R5	R5	4	SD	1	35

^aThe identification number of the study participant.

^bThe name of a unique sequence; the first and second number in the round brackets indicates the number of other sequences with the same V3 in plasma and CSF, respectively, of the participant. The predicted coreceptor usage was determined using Geno2pheno (G2P, with false positive rate of 5%), PhenoSeq and WebPSSM subtype C sini (PSSM-C). Consensus predicted tropism was determined based on the agreement of three, or two of three coreceptor usage prediction algorithms (CPAs). The amino acids occupying positions 11 and 25 (Pos. 11, 25) were identified using PSSM-C; a dash indicates a gap at that position. The number of potential N-linked glycosylation sites (PNGS) in the V3 sequence was counted using N-GlycoSite.

Table 4
The overall predicted coreceptor usage of HIV-1 subtype C in the plasma and CSF of each participants with CM.

ID ^a	Predicted coreceptor usage	
	Plasma	CSF
CM019	R5, R5X4/X4	R5X4/X4
CM021	R5	R5
CM029	n.d	R5
CM032	R5X4/X4	R5X4/X4
CM041	R5	n.d
CM050	R5, R5X4/X4	R5, R5X4/X4
CM052	R5, R5X4/X4	R5
CM054	R5	n.d
CM070	R5, R5X4/X4	n.d
CM089	R5	R5
CM094	n.d	R5
CM098	R5, R5X4/X4	R5, R5X4/X4
CM108	R5, R5X4/X4	R5
CM112	R5	R5
CM117	R5	R5
CM132	R5	R5

^aThe identification number of the study participant. The abbreviation “n.d” (not determined) indicates where were did not generate a sequence for the sample.