Bioinformatic Prediction Programs Underestimate the Frequency of CXCR4 Usage by R5X4 HIV Type 1 in Brain and Other Tissues

Megan E. Mefford[1], **Paul R. Gorry**[2,3], **Kevin Kunstman**[4], **Steven M. Wolinsky**[4], and **Dana Gabuzda**[1,5]

[1] Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Boston, Massachusetts 02115

[2] Macfarlane Burnet Institute for Medical Research and Public Health, Melbourne, Victoria, Australia

[3] Department of Medicine, Monash University, Melbourne, Victoria, Australia

[4] Northwestern University Medical School, Chicago, Illinois 60611

[5] Department of Neurology, Harvard Medical School, Boston, Massachusetts 02115

Abstract

Human immunodeficiency virus (HIV-1) variants in brain primarily use CCR5 for entry into macrophages and microglia, but dual-tropic (R5X4) HIV-1 has been detected in brain and cerebral spinal fluid (CSF) of some patients with HIV-associated dementia (HAD). Here, we sequenced the gp120 coding region of nine full-length dual-tropic (R5X4) *env* genes cloned directly from autopsy brain and spleen tissue from an AIDS patient with severe HAD. We then compiled a dataset of 30 unique clade B R5X4 Env V3 sequences from this subject and 16 additional patients ($n = 4$ brain and 26 lymphoid/blood) and used it to compare the ability of six bioinformatic algorithms to correctly predict CXCR4 usage in R5X4 Envs. Only one program (SVM$_{geno2pheno}$) correctly predicted the ability of R5X4 Envs in this dataset to use CXCR4 with 90% accuracy ($n = 27/30$ predicted to use CXCR4). The PSSM$_{SINSI}$, Random Forest, and SVM$_{genomiac}$ programs and the commonly used charge rule correctly predicted CXCR4 usage with >50% accuracy (22/30, 16/30, 19/30, and 25/30, respectively), while the PSSM$_{X4R5}$ matrix and "11/25" rule correctly predicted CXCR4 usage in <50% of the R5X4 Envs (10/30 and 13/30, respectively). Two positions in the V3 loop (19 and 32) influenced coreceptor usage predictions of nine R5X4 Envs from patient MACS1 and a total of 12 Envs from the dataset (40% of unique V3 sequences). These results demonstrate that most predictive algorithms underestimate the frequency of R5X4 HIV-1 in brain and other tissues. SVM$_{geno2pheno}$ is the most accurate predictor of CXCR4 usage by R5X4 HIV-1.

Human immunodeficiency virus type 1 (HIV-1) infects macrophages and microglia in the central nervous system (CNS) and causes HIV-associated dementia (HAD) or mild neurocognitive impairment in 10–20% of AIDS patients. HIV-1 variants in brain are genetically distinct from those in lymphoid tissues and other organs, and specific sequences in the envelope glycoprotein (Env) coding region of gp160 have been associated with brain compartmentalization.[1] HIV-1 tropism is influenced by the interaction of Env with CD4 and a coreceptor, typically CCR5 or CXCR4. CCR5 (R5) is the primary coreceptor for HIV-1 infection of macrophages and microglia. Several studies identified HIV-1 brain or CSF isolates

capable of mediating entry using CXCR4 (X4 isolates) or both CCR5 and CXCR4 (R5X4 or dual-tropic isolates).[1–4] However, the frequency of X4 or R5X4 strains in the brain of AIDS patients is unknown.

The third hypervariable loop of Env gp120 (V3), a disulfide-linked loop of approximately 35 amino acids, makes direct contact with the coreceptor and is the primary determinant for R5 or X4 tropism.[5] Bioinformatic algorithms that use the V3 sequence to predict HIV-1 coreceptor usage have been developed as a timely and cost-effective alternative to traditional phenotypic assays.[6–10] However, database sequence sets are heavily dependent on R5 sequences and may underestimate X4 usage and dual-tropism. Furthermore, the ability of bioinformatic algorithms to reliably predict X4 usage by R5X4 isolates has not been addressed. Here, we cloned and sequenced nine full-length R5X4 HIV-1 Envs from autopsy brain and spleen tissues from an AIDS patient with severe HAD. We then added these sequences to a larger dataset of unique clade B R5X4 V3 sequences and compared the ability of freely available bioinformatic algorithms to accurately predict X4 usage.

MACS1, a male homosexual patient in the Chicago component of the Multicenter AIDS Cohort Study with no history of antiretroviral therapy, had severe HAD and a $CD4^+$ T cell count of 2 cells/μl at the time of death.[3,4] Analysis of CCR5 alleles by polymerase chain reaction (PCR) demonstrated that the patient was homozygous wild type for CCR5. At autopsy, sections through the frontal and parietal cortex showed pathology consistent with HIV encephalitis (i.e., multiple microscopic foci of necrosis and focal perivascular lesions throughout the white matter occasionally associated with multinucleated giant cells). HIV vacuolar myelopathy and leukoencephalopathy were unusually advanced within the brain stem and cerebellum. We previously isolated four R5X4 HIV-1 viruses from brain (br) and spleen (spln) tissue from this patient [MACS1-br (pbmc), MACS1-br (mdm), MACS1-spln (pbmc), and MACS1-spln (mdm)]. These viruses were isolated from cultures with CD8-depleted peripheral blood mononuclear cells (PBMC) or monocyte-derived macrophages (MDM) as indicated. These R5X4 isolates replicated efficiently in MDM and microglia and induced syncytia formation in >90% of cells by day 10 postinfection. The brain- and spleen-derived isolates entered macrophages and microglia primarily via CXCR4 and induced neuronal apoptosis in primary brain cultures, suggesting that R5X4 variants may be pathogenic in the CNS.[3,4]

To investigate the frequency of R5X4 variants in tissues from patient MACS1, *env* genes were amplified from genomic DNA isolated from autopsy brain and spleen tissues and cloned into the pCR3.1 expression plasmid as described.[4,11] A single-round infection assay screen yielded 28 clones that encoded functional Envs (*n* = 10 brain and 18 spleen clones). Ten Envs from this set (*n* = 5 brain and 5 spleen) were selected for sequencing and further analysis. Expression and processing of 9/10 Envs on 293T cells were verified via Western blotting with antibodies directed against gp120 (goat anti-gp120 from the National Institutes of Health AIDS Research and Reference Reagent Program) (data not shown).[2,4,11] Coreceptor usage was investigated using a cell–cell fusion assay as previously described (Table 1).[4,11] We previously showed that CCR5 and CXCR4 usage determined in this cell–cell fusion assay correlates well with coreceptor usage determined in viral infection assays.[2,3,12,13] The well-characterized ADA (R5), 89.6 (R5X4), and HxB2 (X4) Envs were used as controls. Nine of 10 MACS1 Envs tested (*n* = 5/5 brain and 4/5 spleen) were equally capable of using CCR5 and CXCR4 for fusion in CD4-expressing cells (Table 1). None of the Envs showed a reduced dependence on CD4 levels in the cell–cell fusion assay (data not shown). One Env (spleen-derived clone sp7a-14) was nonfunctional based on cell–cell fusion assays using either CCR5 or CXCR4; this Env contains an amino acid variant (R507) at the gp120/gp41 interface, disrupting the REKR motif required for furin cleavage, which is critical for HIV fusion.[14] Thus, Env sp7a-14 is probably nonfunctional due to loss of gp160 cleavage.

Analysis of amino acid sequences revealed that nine *envs* (four brain and five spleen) encode a full-length gp120 protein. One brain-derived *env* clone (br6b-9) has a short N-terminal truncation with the sequence initiating at the second methionine (position 26) due to a frameshift at position 15 in the N-terminus; this truncation did not affect Env function in the cell–cell fusion assay. Phylogenetic analysis of gp120 nucleotide sequences confirmed distinct compartmentalization of brain- and spleen-derived Env clones (Fig. 1). Phylogenetic analysis of V1V2 amino acid sequences showed tight clustering of brain V1V2 sequences and separation of brain- and spleen-derived Env V1V2 sequences (Fig. 1). MACS1 brain-derived Env clone br6b-8 was more closely related to V1V2 sequences derived from spleen than from brain.

The gp120 V3 loop contains important determinants of coreceptor usage and syncytium induction in MT-2 cells.[5] Consequently, patterns of V3 amino acid variation are frequently used to predict coreceptor usage of primary HIV strains.[6–10,15,16] To determine the ability of freely available bioinformatic algorithms to predict X4 usage in R5X4 Envs, we compiled a dataset containing 30 unique clade B R5X4 V3 sequences (27 unique V3 sequences from 16 patients in 9 published studies and 3 unique V3 sequences from patient MACS1; $n = 4$ brain- and 26 lymphoid/blood-derived sequences) (Fig. 2).[2,12,13,17–22] Envs in this dataset were cloned and sequenced directly from tissue or from low passage isolates, and coreceptor usage of the clones was experimentally determined using viral infection assays[12,17, 20–22] or a combination of viral infection assays and cell–cell fusion assays[2,13] We then used this dataset to determine the ability of six bioinformatic algorithms to accurately predict X4 usage by R5X4 Envs (Fig. 2).

Two simple and commonly used prediction methods are the "11/25" rule and the charge rule. In the "11/25" rule, the presence of a positively charged amino acid at either position 11 or 25 of the V3 loop predicts that the virus can use X4 to mediate entry.[6–9,15,16] This prediction method is significantly less accurate for X4 than for strictly R5-tropic viruses, with <50% and >90% accuracy, respectively.[6,9,16] The charge rule states that an increase in the net charge of the V3 loop (>3) is strongly associated with CXCR4 usage.[10] The position-specific scoring matrix (PSSM) detects nonrandom distributions of V3 amino acids at adjacent sites associated with an empirically determined group of sequences.[8,15] Two separate matrices are available for clade B Envs at http://ubik.microslu.washington.edu/computing/pssm. $PSSM_{X4R5}$ bases predictions on sequences of known coreceptor usage phenotype. $PSSM_{SINSI}$ bases predictions on known syncytium-inducing phenotypes on the MT-2 cell line.[8] Random forest (http://yjxy.ujs.edu.cn/R5-X4pred.rar) evaluates the relative importance of 37 features of the V3 loop including amino acid variation at each position, net charge, and polarity.[10] Finally, two versions of the support vector machine (SVM) algorithm can be used for coreceptor phenotype predictions.[9] $SVM_{genomiac}$ (http://genomiac2.ucsd.edu:8080/wetcat/v3.html) outputs a categorical score (CCR5 or CXCR4) using a dataset aligned to a standard amino acid sequence. $SVM_{geno2pheno}$ (http://coreceptor.bioinf.mpi-sb.mpg.de/cgi-bin/coreceptor.pl) similarly outputs a categorical score based on alignment of V3 nucleotide sequences. The accuracy of these methods in predicting coreceptor usage of R5X4 Envs has not been reported.

Coreceptor usage predictions for the R5X4 V3 sequence dataset using the "11/25" rule were comparable to those reported for X4-tropic V3 sequences, with an accuracy of 43.3% (13/30 Envs). Including position 24 in this rule ("11/24/25" rule) is reported to increase the accuracy of prediction of X4-tropic sequences.[6] However, including position 24 did not affect the prediction accuracy of this R5X4 dataset. The Cardozo *et al.* model was constructed using strictly R5- or X4-tropic Envs. However, R5X4 Envs contain V3 loops that can adopt conformations capable of interacting with either CCR5 and CXCR4, and may therefore contain surface patches that do not conform to the static models of the V3 structure in this proposed model.[6] A second commonly used coreceptor-prediction method, the charge rule, predicted

that 83.3% of the V3 data set could use CXCR4 for entry (25/30 Envs). Of the two PSSM matrices, the $PSSM_{X4R5}$ matrix correctly predicted X4 usage for 10/30 V3 sequences (33.3% accuracy). $PSSM_{X4R5}$ predicted that 3/3 MACS1 unique V3 sequences would be strictly R5-tropic (0% accuracy). The $PSSM_{SINSI}$ matrix correctly predicted X4 usage in 22/30 V3 sequences, thereby increasing the accuracy of prediction for our dataset of R5X4 Envs from 33.3% to 73.3%. X4 usage was correctly predicted for 2/3 unique MACS1 V3 sequences (66.6% accuracy). The random forest program predicted X4 usage for 16/30 Envs, with an accuracy of 53.3%, and predicted that 3/3 unique MACS1 Envs were strictly R5-tropic (0% accuracy). $SVM_{genomiac}$ correctly predicted that 19/30 Envs could use CXCR4 to mediate entry, with an accuracy of 63.3%. As with $PSSM_{SINSI}$, X4 usage was correctly predicted for 2/3 unique MACS1 V3 sequences (66.6% accuracy). Finally, $SVM_{geno2pheno}$ predicted that 27/30 R5X4 Envs could use X4 for entry, with an accuracy of 90.0% (specificity, defined here as the rate of false positives, was set at 10%). Prediction of X4 usage for the unique MACS1 Envs was 3/3 (100%). X4 usage predictions were most accurate when specificity levels were at 5–10%. Increasing the stringency of predictions by decreasing the specificity rate to <5% resulted in a concurrent decrease in prediction accuracy, but could be overcome by inputting clinical data (CCR5 genotype and $CD4^+$ counts).

While no single variant appears to affect the accuracy of coreceptor prediction in every genetic background, changing the Env V3 sequence entered into bioinformatic programs to assess the effects of specific amino acids on prediction accuracy in our R5X4 dataset identified several positions that may contribute to the accuracy of coreceptor prediction using bioinformatic approaches. Both the X4R5 and SINSI PSSM matrices incorrectly predict that MACS1 brain Env clone br6b-8 is R5-tropic. Env br6b-8 contains the consensus alanine at position 19 in the V3 loop, while the other eight MACS1 Env clones contain a valine at this position. Changing the amino acid sequence from our dataset of 30 unique V3 sequences from valine to alanine at position 19 also changes the PSSM prediction from X4- to R5-tropic in nine additional Envs. These results suggest that an alanine-to-valine change at position 19 may be associated with a change from R5- to X4-tropism in sequences from the training sets of both matrices. Another interesting finding from the PSSM matrix predictions is that a glutamine-to-lysine variant at position 32 in the V3 loop changed the prediction for our set of MACS-1 Env clones from R5-tropic to X4-tropic. This trend is applicable to additional V3 sequences from our dataset that contain the lysine variant at position 32. These results suggest that inclusion of V3 amino acid sequences from more X4 or R5X4 Envs to algorithm training sets will further increase the prediction accuracy of these matrices.

In summary, we identified R5X4 HIV-1 in brain from a patient with severe HAD. Comparison of the ability of bioinformatic algorithms to correctly predict X4 usage by R5X4 Envs in our dataset showed that the frequency of R5X4 HIV-1 is underestimated by most commonly used predictive algorithms. $SVM_{geno2pheno}$ is the most accurate predictor of CXCR4 usage by R5X4 HIV-1 in brain and other tissues. Bioinformatic prediction tools provide a convenient method to screen for coreceptor usage, an issue of increasing importance for clinicians considering the use of CCR5 antagonists in HIV-infected patients. It will therefore be important for future studies to increase X4 and R5X4 sequences associated with bioinformatic algorithm training set sequences in the development of prediction tools in order to better define the patterns of amino acid variation that contribute to inaccurate predictions.

Acknowledgments

References

1. Dunfee R, Thomas ER, Gorry PR, et al. Mechanisms of HIV-1 neurotropism. Curr HIV Res 2006;4:267–278. [PubMed: 16842080]

2. 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;77:12336–12345. [PubMed: 14581570]

3. 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;75:10073–10089. [PubMed: 11581376]

4. Gorry PR, Taylor J, Holm GH, et al. Increased CCR5 affinity and reduced CCR5/CD4 dependence of a neurovirulent primary human immunodeficiency virus type 1 isolate. J Virol 2002;76:6277–6292. [PubMed: 12021361]

5. Hoffman TL, Doms RW. HIV-1 envelope determinants for cell tropism and chemokine receptor use. Mol Membr Biol 1999;16:57–65. [PubMed: 10332738]

6. Cardozo T, Kimura T, Philpott S, et al. Structural basis for coreceptor selectivity by the HIV type 1 V3 loop. AIDS Res Hum Retroviruses 2007;23:415–426. [PubMed: 17411375]

7. Resch W, Hoffman N, Swanstrom R. Improved success of phenotype prediction of the human immunodeficiency virus type 1 from envelope variable loop 3 sequence using neural networks. Virology 2001;288:51–62. [PubMed: 11543657]

8. 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;77:13376–13388. [PubMed: 14645592]

9. Low AJ, Dong W, Chan D, et al. Current V3 genotyping algorithms are inadequate for predicting X4 co-receptor usage in clinical isolates. AIDS 2007;21:F17–24. [PubMed: 17721088]

10. Xu S, Huang X, Xu H, et al. Improved prediction of coreceptor usage and phenotype of HIV-1 based on combined features of V3 loop sequence using random forest. J Microbiol 2007;45:441–446. [PubMed: 17978804]

11. Thomas ER, Dunfee RL, Stanton J, et al. Macrophage entry mediated by HIV Envs from brain and lymphoid tissues is determined by the capacity to use low CD4 levels and overall efficiency of fusion. Virology 2007;360:105–119. [PubMed: 17084877]

12. Gray L, Churchill MJ, Keane N, et al. Genetic and functional analysis of R5X4 human immunodeficiency virus type 1 envelope glycoproteins derived from two individuals homozygous for the CCR5delta32 allele. J Virol 2006;80:3684–3691. [PubMed: 16537640]

13. Gorry PR, Dunfee RL, Mefford ME, et al. Changes in the V3 region of gp120 contribute to unusually broad coreceptor usage of an HIV-1 isolate from a CCR5 delta32 heterozygote. Virology 2007;362:163–178. [PubMed: 17239419]

14. Bosh VMP. Mutational analysis of the human immunodeficiency virus type I env gene product proteolytic cleavage site. J Virol 1990;64:2337–2344. [PubMed: 2182911]

15. Shah M, Smit TK, Morgello S, et al. Env gp120 sequence analysis of HIV type 1 strains from diverse areas of the brain shows preponderance of CCR5 usage. AIDS Res Hum Retroviruses 2006;22:177–181. [PubMed: 16478400]

16. Hoffman NG, Seillier-Moiseiwitsch F, Ahn J, et al. Variability in the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-associated changes in the V3 loop. J Virol 2002;76:3852–3864. [PubMed: 11907225]

17. Yi Y, Shaheen F, Collman RG. Preferential use of CXCR4 by R5X4 human immunodeficiency virus type 1 isolates for infection of primary lymphocytes. J Virol 2005;79:1480–1486. [PubMed: 15650174]

18. 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;70:1651–1667. [PubMed: 8627686]

19. Sullivan PS, Schable C, Koch W, et al. Persistently negative HIV-1 antibody enzyme immunoassay screening results for patients with HIV-1 infection and AIDS: Serologic, clinical, and virologic results. Seronegative AIDS Clinical Study Group. AIDS 1999;13:89–96. [PubMed: 10207549]

20. Skrabal K, Saragosti S, Labernardiere JL, et al. Human immunodeficiency virus type 1 variants isolated from single plasma samples display a wide spectrum of neutralization sensitivity. J Virol 2005;79:11848–11857. [PubMed: 16140761]

21. van Rij RP, Blaak H, Visser JA, et al. Differential coreceptor expression allows for independent evolution of non-syncytium-inducing and syncytium-inducing HIV-1. J Clin Invest 2000;106:1569. [PubMed: 11120764]

22. Hu QX, Barry AP, Wang ZX, et al. Evolution of the human immunodeficiency virus type 1 envelope during infection reveals molecular corollaries of specificity for coreceptor utilization and AIDS pathogenesis. J Virol 2000;74:11858–11872. [PubMed: 11090186]
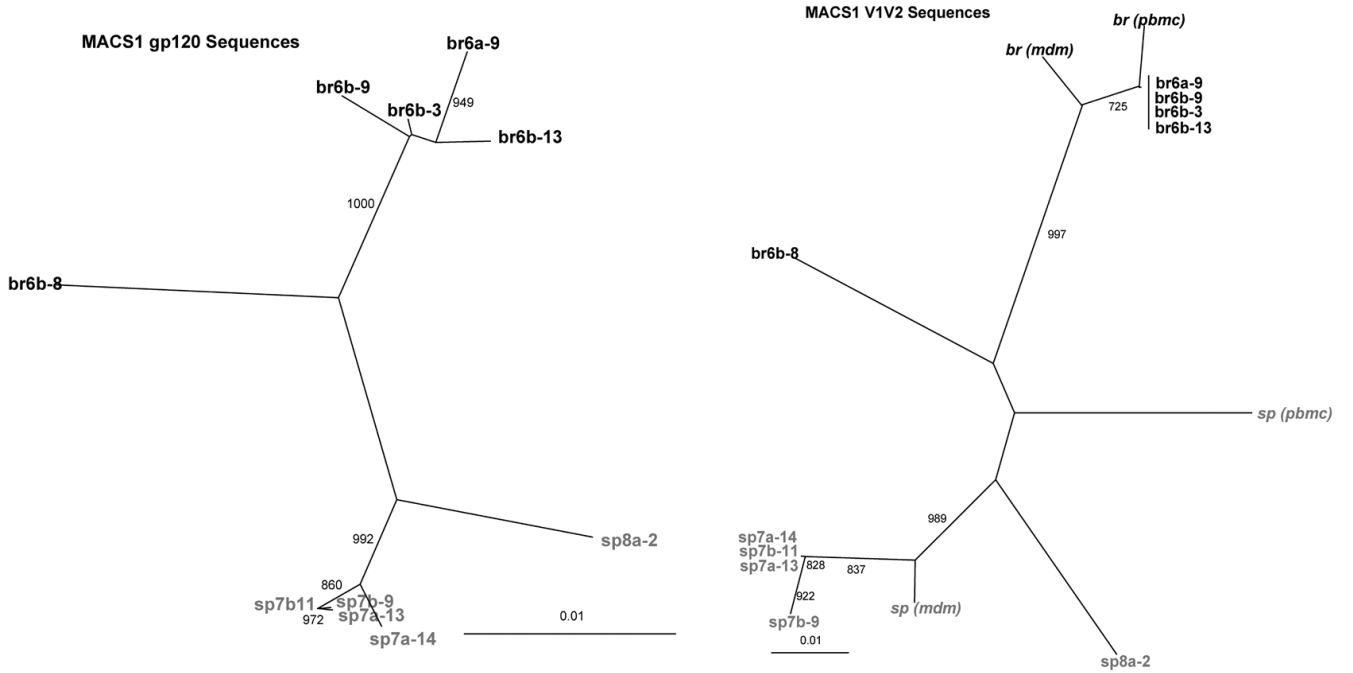
FIG. 1.

Phylogenetic analysis of HIV-1 gp120 and V1V2 sequences. Sequences from brain and spleen tissues are color coded black and gray, respectively. Numbers associated with each branch are bootstrap values, which represent the number of trees, out of 1000 replicates performed, in which the same branching order was found. Only values above 800 for the major branches are shown. Branch lengths are proportional to the amount of sequence divergence. Scale bars indicate 1% sequence divergence. The left panel depicts the phylogenetic relationship among gp120 nucleotide sequences amplified directly from brain and spleen tissues from patient MACS1. The right panel depicts V1V2 amino acid sequences derived from patient MACS1. MACS1 brain and spleen viral isolates obtained by PBMC coculture are shown in italics.[3,4]

A

```
Clade B         CTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHC
MACS1br6b-8     ...........MSL...K......Q.......K...(2,7)
MACS1br6b-9     .I.........MSL...KV.....Q......K...(2,4,6,7)
MACS1br&sp      ...........MSL...KV.....Q......K...(2,4,6,7)
aBR01           .........R.VTM.........D............
aBL01           ........KR..TR....VY....D.V.........(2,3,4,6,7)
dBR02/07        .........RKGSM......LAREQ...N.......(1,2,3,4,5,6,7)
15888           .........RG.Y......I...DK.V....K.Y.(1,2,3,4,5,6,7)
C2-3/16         ...........TM....VY.............(7)
C2-22           ...........TM....VY...I.........(4,7)
C2-24           ...........TM....VD.............(7)
DR-1            .I.........MTL...KV.....-VT.S..K...(2,4,5,6,7)
DR-8/17         .I.........MTL...KV.....-VT....K...(2,4,6,7)
DR-19           .I.........MTL...KV.....-VT....KE..(2,4,6,7)
89.6            .........RRLS........ARRN..........(1,2,3,4,5,6,7)
DH12            .........G.TL...V.......V....K...(2,4,7)
T5-R5X4-1/2     .........R.TM....VY............K...(1,2,3,4,5,6,7)
3921F12+3       .........R.SLS...V............K...(1,2,3,4,5,6,7)
20813B1         .........RR.Y..Q...V...RQ......K.Y.(1,2,3,4,5,6,7)
92HT594         .........R.S......S....Q......K...(1,2,3,4,5,6,7)
17141raf5       ....G....RR.S..........EQ...N.......(1,2,3,4,5,6,7)
17136roe11      ....G.....R.S..........EQ...N.......(1,2,3,4,5,6,7)
17131rog6       ....G.....R...........EQ...N......(1,2,4,5,6,7)
17129roa10      ....G.....R...........RQ...N......(1,2,4,5,6,7)
17136rod11      ....G.....R.S.........EQ..........(1,2,4,5,6,7)
QH1520c.2       .............RM.IR.....RE-..........(2,5,7)
QH1549c.13      .........I...RM.IR.....RE-..........(2,4,5,7)
96USHIPS9       .........G.N.....TV.A..K..........(1,2,4,5,7)
49030rog6       .........G..L.....L.A..G.V........(2)
409028rod7      ............L..........G...N......(2)
QH1521c.34      .........H..L.A.K....GE-..........(6,7)
```

B

Predictive Algorithm	Correct Predictions/ Total # R5X4 Envs	Accuracy
11/24/25 Rule	13/30	43.3%
Charge Rule	25/30	83.3%
$PSSM_{X4R5}$	10/30	33.3%
$PSSM_{SINSI}$	22/30	73.3%
Random Forest	16/30	53.3%
$SVM_{genomiac}$	19/30	63.3%
$SVM_{geno2pheno}$	27/30	90.0%

FIG. 2.

Prediction of coreceptor usage by R5X4 V3 sequences using bioinformatic prediction algorithms. (**A**) Unique R5X4 V3 amino acid sequences ($n = 30$) from 17 patients were aligned using ClustalX. The sequences are shown aligned to the CCR5-tropic Clade B consensus sequence. Sequences are grouped by patient. Sequences derived from brain ($n = 4$) are underlined. Positions 11 and 25 in the V3 loop are highlighted in the sequence alignments. Algorithms that correctly predict the ability to use CXCR4 are indicated following each sequence. 1 = 11/24/25 rule; 2 = charge rule; 3 = PSSM (X4R5); 4 = PSSM (SINSI); 5 = random forest; 6 = $SVM_{genomiac}$; 7 = $SVM_{geno2pheno}$. (**B**) Summary of total number of correct

predictions for each bioinformatic algorithm and the program's specificity for the 30 R5X4 V3 sequences.

Table 1

MACS1 Envs Use Both CCR5 and CXCR4 to Mediate Fusion[a]

Envelope	CCR5	CXCR4
ADA	++++	−
89.6	+++	++++
HXB2	−	++++
MACS1br6b-8[1]	++	++
MACS1br6b-9[1]	++	++
MACS1br6b-3	+++	+++
MACS1br6b-13[2]	++++	+++
MACS1br6a-9	++++	+++
MACS1sp7b-11	+++	++++
MACS1sp8a-2	+++	++++
MACS1sp7a-14[2]	−	−
MACS1sp7a-13	+++	++++
MACS1sp7b-9	+++	++++

[a]293T cells cotransfected with pCR3.1Env and pLTR-Tat were mixed with Cf2-Luc cells cotransfected with pcDNA3-CD4 and either pcDNA3-CCR5 or pcDNA3-CXCR4. Cells were harvested following an 8-h incubation. The levels of fusion in cells expressing CD4 and CCR5 or CXCR4 as measured by luciferase activity are indicated as follows: −, +, ++, +++, and ++++ representing levels that were at background, 1–25%, 25–50%, 50–75%, and 75–100% of control Envs, respectively. Unless otherwise indicated, results are representative of five independent assays.

[1]Results from one fusion assay.

[2]Results from two fusion assays.