International Journal of Advanced Biotechnology and Research
ISSN 0976-2612, Online ISSN 2278–599X,
Vol 4, Issue 1, 2013, pp 105-112
http://www.bipublication.com

GENETIC CHARACTERISTICS OF THE HYPER-VARIABLE REGION (V3-V5) OF *env* GENE OF HIV-1 DISTINGUISHING SUBTYPE C FROM SUBTYPE B

Supriya Singh[1, 2], Shashi Khare[1], D S Rawat[1], Saaransh Jain[1], R L Ichhpujani[1], Sudha Prasad[2], Megha Singhal[1,2], L S Chauhan[1] and Arvind Rai[1*]

[1]National Centre for Disease Control (NCDC), Delhi, India
[2]Maulana Azad Medical College (MAMC), India
*Corresponding author: Email: drarvindrai21@gmail.com, Phone No.: +91 9810385357

ABSTRACT

Aim: Carry out molecular characterization of subtype B and C sequences encompassing the hyper-variable region (V3-V5) of envelope (env) gene of HIV-1 to determine the genetic characters that distinguish and may possibly account for differential pathogenicity of the two subtypes. **Experimental Design:** Blood samples were collected from 25 HIV-1 sero-positive patients from India. Hyper-variable region (V3-V5) of env gene was amplified and sequenced. 173 sequences belonging to subtypes B and C were retrieved from Los Alamos database. Co-receptor prediction, antigenic loop, co-receptor variability in CCR5 and CXCR4 utilizing subtype B and C sequences, N glycosylation site, signature sequence and phylogenetic analysis was carried out. **Results:** Loss of N glycosylation sites in CXCR4 co-receptor utilizing sequences were observed in 37% of subtype B and 23% of subtype C sequences. Antigenic tip in majority of the subtype C isolates was GPGQTFY. Significant entropy differences were observed in CCR5 and CXCR4 co-receptor utilizing sequences of subtypes B and C. We observed several signature amino acid residues which occurred in high frequencies in Subtype C CCR5/CXCR4) co-receptor utilizing sequences. **Conclusions:** In conclusion, the determined molecular characteristics of subtypes B and C may partially be involved in the high pathogenicity of subtype C.

Keywords: HIV; Co-receptor; V3 loop; N-glycosylation; Envelope; Hyper-variable

INTRODUCTION

As per the UNAIDS Report 2010, since the beginning of the epidemic of HIV in 1981 the deadly virus has infected almost 60 million people worldwide and claimed about 25 million lives due to HIV-related causes (1). India harbours world's third largest HIV burden after South Africa and Nigeria. Estimates of the department of AIDS control (2009-10) report that 2.27 million people are living with HIV/AIDS in India (2). Human immunodeficiency virus type-1 (HIV-1) exhibits an extensive genetic variability. Among the different subtypes of HIV-1, subtypes B and C are the major subtypes worldwide. HIV-1 subtype C is the predominant subtype in India and is considered to be more pathogenic than subtype B (3, 4, 5, 6). The envelope (*env)* gene (gp120) plays a crucial role in the process of entry of HIV-1 into target cells. The five variable regions (V1-V5) of the *env* encompass the potential pathogenic region of the virus. The V1-V2 region influence replication efficiency, V3 region is important in host cell tropism, V4 and V5 regions are involved in CD4 binding and neutralizing antibody responses (7). The virus attacks human cells by interacting with CD4 receptor which is followed by a conformational change in gp 120. This exposes the co-receptor binding site for CXCR4 or/and CCR5 (major co-receptors for HIV).

Based on differences in growth properties and cytopathic effects on PBMC's, HIV produces syncytium inducing (SI) and non-syncytium inducing (NSI) viruses in vitro. NSI viruses utilize CCR5 as their major co-receptor, whereas SI viruses utilize the CXCR4 co-receptor. Newly infected patients primarily exhibit CCR5 (R5) co-receptor usage whereas patients in later stages exhibit CXCR4 (X4) co-receptor usage (8). Co-receptor determination can be useful in monitoring disease progression. V3 loop of HIV is a significant region of the hyper-variable region of *env* as its amino acid sequence can be used for prediction of co-receptor and give a fair idea about the utilized co-receptor while avoiding expensive phenotypic tests. N-glycosylation of the V3 loop shields the virus and prevents it from the neutralizing antibody responses/ host immune responses (9). The antigenic tip of the V3 loop possesses a conserved secondary structure. It is also a target for hosts neutralizing antibodies and is therefore an important determinant of viral pathogenicity (10). Most of the published data on co-receptor utilization, replication efficiency and cytopathic effect relates to HIV- 1 subtype B; while limited information is available with regard to subtype C (11, 12, 13, 14). The present study was carried out on Indian samples and sequences retrieved from database to understand the epidemic of HIV-1 subtypes B and C through phylogenetic analysis and molecular characterization of the two subtypes.

MATERIALS & METHODS

Sample collection and study subjects

Ethical approval was obtained from the ethical committee of National Centre for Disease Control (NCDC) and Maulana Azad Medical College (MAMC) & Associated Lok Nayak Hospital. HIV positive cases from different risk groups including heterosexual males, heterosexual females and antenatal risk group were recruited for the study. Informed consent was obtained from all participants.

Peripheral Blood Mononuclear Cells (PBMC) Isolation and DNA Extraction

Whole blood (3-4ml) was drawn in K2EDTA vacutainers and PBMC's were isolated immediately after collection using Ficoll's Reagent. Genomic DNA, including proviral DNA of HIV was extracted with the help of QIAamp DNA Blood Mini kit, according to the manufacturer's protocol.

Polymerase Chain Reaction (PCR)

PCR for *β-globin* was carried out as described previously by Khoja S et al (15). This was done to check the quality of the genomic DNA. PCR was carried out for partial *gag* gene of HIV to confirm the serological diagnosis. PCR was carried out as has been described previously by Grankvist et al. PCR for partial *env* gene (V3-V5) of HIV was carried out (11). Amplifications were carried out in ABI 9700 thermal cycler. The PCR reaction was carried out using Go Taq PCR core system II (Promega) as per the kit protocol.

Automated nucleotide sequencing

PCR products were purified using Promega Wizard SV Gel & PCR Clean up system. Purified PCR products were subjected to automated nucleotide sequencing with forward and reverse primers separately. Sequencing was carried-out using Big Dye terminator kit (ABI, USA) using the kit protocol.

Analysis of Sequences

Gene sequences for *env* obtained in the present study were submitted to GenBank at www.ncbi.nlm.nih.gov and accession numbers were obtained. A BLAST search was carried out in order to confirm identity of the strains. Samples were assigned subtypes based on closest homology found with the subtype references in NCBI. DNA and protein alignments were created using Clustal X program. REGA subtyping tool was used to determine the subtype of our samples (16, 17). Phylogenetic analysis was performed using maximum likelihood with the help of MEGA version 5 and the reliability of the branching orders was determined by bootstrap method (18).

Co-receptor prediction: Co-receptor prediction was performed using the PSSM tool (C-PSSM, *http://indra.mullins.microbiol.washington.edu/webpssm/*), (ds) Kernel (*http://genome.ulaval.ca/hiv-dskernel*), Geno2pheno (*http://coreceptor.bioinf.mpi-inf.mpg.de/*).

Antigenic tip: Deduced amino acid sequences from the nucleotide sequences, using MEGA software, was used to study antigenic tip.

Sequence variability: Sequence variability was determined in the V3 loop for CCR5/CXCR4 co-receptor utilizing subtypes B & C sequences using "Entropy two tool" in Los Alamos Database. This tool was used to construct sets of aligned sequences and determine any positions with significant variation at the corresponding position in the other set (18).

For the sequence variability study we retrieved 51 sequences from CCR5 utilizing subtype B sequences [AF310109-AF310115, AF310124-AF310127, AF310131, AF491737-AF491742, AF540994-AF541001, AF541102-AF541114, AJ418478-AJ418484, AJ4184850, AJ418486-AJ418488], 43 CCR5 utilizing sequences [EU161644-EU161645, EU293445-EU293450, EU786673, EU786681, EU855133, EU884500, FJ653278- FJ653294, HM60277-HM60279, HM60281-84, HM60285-HM60292], 36 CXCR4 utilizing subtype B sequences [A04321,AF033819, AF035534, AF07572, AF075721, AF075722, AF146728, AF189159, AF310132, AJ810482-AJ810484, AY173951, AY173956, AY189526, AY736819, AY736820, AY842799, AY842819, AY842825,AY842826, AY842828, AY842831-AY842834, DQ177202, DQ286957, EU578431, EU604587, FJ653148,K03455, L31963, M14100, U08447, X01762] and 43 CXCR4 utilizing sequences subtype C sequences[AF411966, AY265948, AY265949, AY529666, AY529677-AY529679, DQ382362, DQ382372, DQ382378,DQ904342, DQ904343, FJ846628-FJ846637, FJ846640, FJ846641, FJ846643-FJ846660, L22956].

RESULTS

25 HIV-1 sero-positive subjects were enrolled in the study from various risk groups, 117 including heterosexual male, heterosexual female and antenatal risk group.

Amplification of samples by PCR

All 25 samples were found to be positive for the ~142 bp region of *gag* gene of HIV. The samples were then amplified for hyper-variable (V3-V5) region of *env* gene of HIV using nested PCR. Only 16 samples exhibited amplification for the *env* gene.

Sequencing and GenBank Submission

The sequences obtained were submitted to GenBank and accession numbers were obtained (HM630277- HM630292).

BLAST search and phylogenetic analysis

BLAST search revealed that our isolates belonged to HIV-1 *env* gene (V3-V5). The samples were found to belong to subtype C. Samples were aligned using Bioedit vesion 7.0.9.0. Phylogenetic analysis was carried out using the MEGA software version 5 with Maximum likelihood method (Figure 1).

N-linked glycosylation

In the CXCR4 utilizing V3 sequences, loss of the glycosylation site was observed in 37% sequences in subtype B sequences and 23% sequences in subtype C sequences.

Antigenic tip

Central domain of the V loop identified as the principal neutralizing determinant comprising GPGR motif determined the cross clade neutralizing activity of a human monoclonal antibody. Our isolates showed GPGQTFY to be the antigenic tip most common in subtype C V3 loop sequences in about 87.5% cases. GPGQAFY motif commonly found in subtype B isolates was found in 12.5% of our subtype C isolates.

Co-receptor prediction

Fourteen sequences were found to be predictive of utilizing CCR5 co-receptor by the three methods.

V3 loop sequence variation

In both subtypes B and C, increase in sequence variability was observed in the V3 loop sequences in CXCR4 co-receptor utilizing sequences when compared with CCR5 co-receptor utilizing sequences (Figure 2). The subtype C exhibited increased sequence variability for CCR5 and CXCR4 co-receptor utilizing sequences when compared with subtype B.

Signature Residues

Signature residues include arginine at position 15 (in 96.6% isolates), glutamine 146 at position 23 (in 100% isolates) in the CCR5 co-receptor utilizing

sequences in subtype C. A signature residue in the CXCR4 co148 receptor utilizing sequence was at position 15 (in 86%) in subtype C. [Figure 3, Table 2].

DISCUSSION

Out of the 25 HIV-1 *gag* gene positive samples, only 16 (~64%) successfully amplified the V3-V5 region of the *env* gene. The unsuccessful amplification in the 9 samples could be due to mutations in the primer binding region in our Indian isolates. All sequences from this study clustered with previously published subtype C sequences indicating that subtype C is the subtype currently circulating in the Indian isolates. Regarding the epidemiology and transmission of the epidemic of HIV, with reference to subtypes B and C, sequences of subtype B, subtype C and recombinants clustered separately with the bootstrap values of the major nodes being very high. The neighbouring countries of India such as China and Myanmar exhibited clustering of isolates with our sequences. This is indicative of transmission of the infection between India, China and Myanmar and thus is essential to monitor the risk behaviours through surveillance in order to prevent/reduce transmission. N-glycosylation of variable loops, such as the V3 loop, often restricts access to conserved host receptor binding sites. This limits their exposure to the host immune system, and confers a survival advantage for the virus. We observed in the CXCR4 utilizing V3 sequences loss of glycosylation site 37% sequences in subtype B sequences and loss of 23% glycosylation sites in subtype C sequences when compared with their respective glycosylation sites in CCR5 sequences. The loss of glycosylation sites in CXCR4 sequences suggests that possibly in the later stages of infection the loss of glycosylation sites providing a survival advantage to the virus. This advantage is observed in a higher percentage of B subtype B as compared with subtype C. This possibly may be due to the fact that the virus in subtype B mainly utilizes both co-receptors (CCR5 in early stages and CXCR4 in late stages of HIV infection), whereas in subtype C the virus utilizes mostly CCR5 co-receptor even in late stages. Thus, this viral escape strategy is not of much significance in subtype C, but is certainly important for the B subtype. The antigenic tip in the subtype C isolates from our study reveals the predominant antigenic tip to be GPGQTFY in 87.5% of our isolates, the rest showed GPGQAFY. Antigenic tip analysis is important for epitope-specific neutralizing antibody preparation and vaccine development. As has already been reported, the numbers of CXCR4-using isolates in subtype C are limited, and so was the case in our study, where all isolates were predictive of utilizing CCR5 co-receptor. It is possible that there might be a viral adaptation within the subtype C V3 loop that allows for a limited CXCR4 co-receptor usage. Signature residues in the V3 loop were therefore studied. The signature residues possibly could account for a high pathogenicity, since these occur in high percentages in subtype C CCR5/CXCR4 co-receptor utilizing isolates and the corresponding residue was replaced in subtype B. Effect of these amino acid residues needs to be studied in order to better understand the subtype's high pathogenicity.

During the initial asymptomatic stage of infection or the CCR5 co-receptor utilizing stages, the sequence variability is less as the immune system exerts a selective pressure on the virus and therefore its sequence remains comparatively conserved. Eventually, as the infection progresses towards symptomatic phase or towards CCR4 utilizing co-receptor stages, the sequence variability increases as immune system of the body is now immune-compromised and is not able to exert selective pressure and allows the virus to become more variable. When comparing the CCR5 utilizing sequences in subtype B and C, significantly increased variability in the subtype C and similar is the case with CXCR4 utilizing co-receptor sequences of subtype C. The subtype C HIV-1 virus is infectious and the body is not so efficient in exerting a selective pressure on the virus, thus leading to comparatively increased entropy.

In conclusion, our results on molecular characteristics of subtypes B and C show that these

possibly may be involved in the subtype C's high pathogenicity.

Funding

Partial funding was received from Council of Scientific Research

Conflict of Interest

Authors declare no conflict of interest in this manuscript.

Ethical Approval Declaration

Ethical approval was obtained from the ethical committee of Maulana Azad Medical College & National Centre for Disease Control, Delhi.

ACKNOWLEDGEMENTS

Author SS acknowledges financial support from the Council of Scientific and Industrial Research (CSIR), Delhi, India, during the course of this study. We thank SS Negi, RS Rautela, Hanu Ram, Poonam Gupta and Sachin Kumar for helpful discussions. Authors also thank the staff of the HIV/AIDS Division of NCDC and for their assistance in collecting samples, and providing serological and clinical data.

REFERENCES

1. UNAIDS Report on the Global AIDS Epidemic 2010, available from URL
http://www.unaids.org/globalreport/Global_report.htm.
2. Department of AIDS Control, Ministry of Health & Family Welfare, Annual Report 2009, available from URL http://www.nacoonline.org/upload/AR%202009-10/NACO_AR_English%20corrected.pdf.
3. Mandal D, Jana S, Bhattacharya SK, Chakrabarti S. (2002) HIV type 1 subtypes circulating in eastern and northeastern regions of India. AIDS Res Hum Retroviruses 18:1219-27.
4. Shankrappa R, Chatterjee R, Learn GH, Neogi D, Ding M, Roy P, Ghosh A, Kingsley L, Harrison L, Mullins JI, Gupta P. (2001) Human immunodeficiency virus type I *env* sequences from Calcutta in eastern India: Identification of features that distinguish subtype C sequences in India from other subtype C sequences. J Virolo. ; 75:10479-10487.
5. Jameel S, Zafrullah M, Ahmad M, Kapoor GS, Sehgal S. (1995) A genetic analysis of HIV-1 from Punjab, India reveals the presence of multiple variants. AIDS. 9: 685-690.
6. Vasudha Sundarvaradan, Suman R Das, Rajesh Ramakrishnan, Shobha Sehghal, Sarla Gopalan, Nafes Ahmad, Shahid Jameel. (2007) Role of HIV-1 subtype C envelope V3 to V5 regions in viral entry, co-receptor utilization and replication efficiency in primary T-lymphocytes and monocyte-derived macrophages. Virology Journal. 4:126.
7. Sirois S, Sing T, Chou KC.(2005) HIV-1 gp120 V3 Loop for Structure-Based Drug Design. Current Protein and Peptide Science. 6: 413-422.
8. Maas JJ, Gange SJ, Schuitemaker H. (2000) Strong association between failure of T cell homeostasis and the syncytium-inducing phenotype among HIV-1-infected men in the Amsterdam Cohort Study. AIDS 14:1155-61.
9. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA. (2001) N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. J Biol Chem 276:13433-41.
10. Dong XN, Wu Y, Ying J, Chen YH. (2005) The antigenic tip GPGRAFY of the V3 loop on HIV-1 gp120: genetic variability and subtypes. Immunol Lett 101:112-4.
11. Delwart EL, Shaper EG, Louwqagie J, et al. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. Science 1993; 262: 1257-61
12. Connell BJ, Michler K, Capovilla A, Venter WD, Stevens WS, Papathanasopoulos MA. (2008) Emergence of X4 usage among HIV-1 subtype C: evidence for an evolving epidemic in South Africa. AIDS 22: 896-9.
13. Lynch RM, Shen T, Gnanakaran S, Derdeyn CA. (2009) Appreciating HIV type 1 diversity: subtype differences in Env. AIDS Res Hum Retroviruses 25: 237-248.
14. Coetzer M, Nedellec R, Cilliers T, Meyers T, Morris L, Mosier DE. (2011) Extreme genetic divergence is required for coreceptor switching in HIV-1 subtype C. J Acquir Immune Defic Syndr 56:9-15.
15. Khoja S, Ojwang P, Khan S, Okinda N, Harania R, Ali S. (2008) Genetic Analysis of HIV-1 Subtypes in Nairobi, Kenya. PLoS one 3:e3191.
16. Alcantara LCJ, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, Galvao-Castro B, Vandamme A-M, and de Oliveira T. (2009) A Standardized Framework for Accurate, High-throughput Genotyping of Recombinant and Non-recombinant Viral Sequences. Nucleic Acids Research: 37(Web Server issue):W634-42. Epub.
17. Oliveira T, Deforche K, Cassol S, Salminem M, Paraskevis D, Seebregts C, Snoeck J, van Rensburg

EJ, Wensing AMJ, van de Vijver DA, Boucher CA, Camacho R, and Vandamme A-M. (2005) An Automated Genotyping System for Analysis of HIV-1 and other Microbial Sequences. Bioinformatics 21: 3797-3800.

18. Efron B and R Tibshirani. (1991) Statistical Data Analysis in the Computer 287 Age. Science 253: 290-395.

Table 1: Table showing the risk group, predicted co-receptor, and clinical stage of infection.

Accesssion Number	Risk group	Age/Sex	Web PSSM Score	Web PSSM Tropism	Geno2pheno	(ds) Kernel	Clinical manifestations
HM630277	ANC	19/F	-30.51	R5	R5	R5	Asymptomatic
HM630278	ANC	27/F	-29.93	R5	R5	R5	Asymptomatic
HM630279	HM	29/M	-29.93	R5	R5	R5	Asymptomatic
HM630280	HF	35/F	-24.05	R5	X4	R5	Asymptomatic
HM630281	ANC	25/F	-28.60	R5	R5	R5	Asymptomatic
HM630282	HM	29/M	-29.07	R5	R5	R5	Asymptomatic
HM630283	HM	30/M	-29.13	R5	R5	R5	Asymptomatic
HM630284	ANC	22/F	-29.93	R5	R5	R5	Asymptomatic
HM630285	ANC	23/F	-18.91	X4	R5	R5	Asymptomatic
HM630286	HF	28/F	-26.17	R5	R5	R5	Asymptomatic
HM630287	HM	37/M	-29.93	R5	R5	R5	Asymptomatic
HM630288	ANC	20/F	-29.62	R5	R5	R5	Asymptomatic
HM630289	ANC	24/F	-29.17	R5	R5	R5	Asymptomatic
HM630290	HF	29/F	-29.05	R5	R5	R5	Asymptomatic
HM630291	HF	31/F	-28.47	R5	R5	R5	Asymptomatic
HM630292	ANC	22/F	-26.17	R5	R5	R5	Asymptomatic

*ANC: Antenatal care; HM: Heterosexual male; HF: Heterosexual female

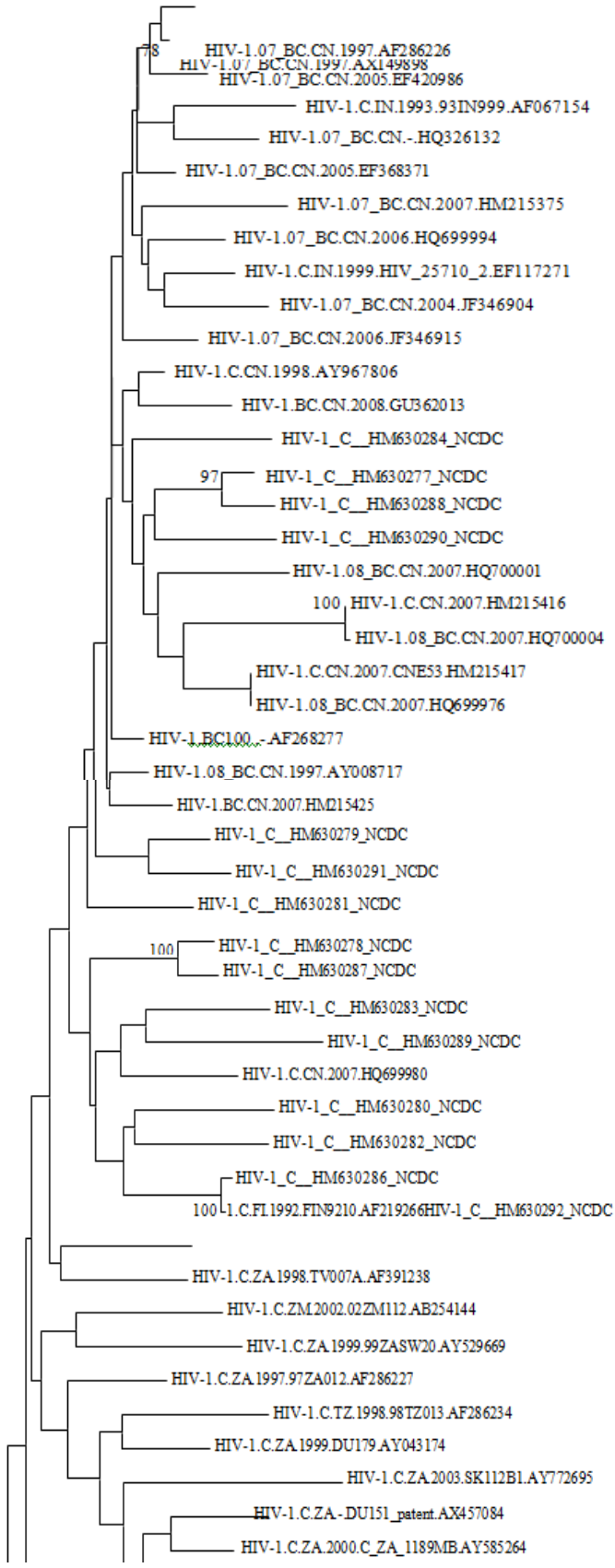
Table 2: (a) Table showing signature residues in the CCR5 co-receptor utilizing sequences of V3 loop. (b) Table showing the signature residues in the CXCR4 co-receptor utilizing sequences of V3 loop (Subtype C-Query; Subtype B-Background)

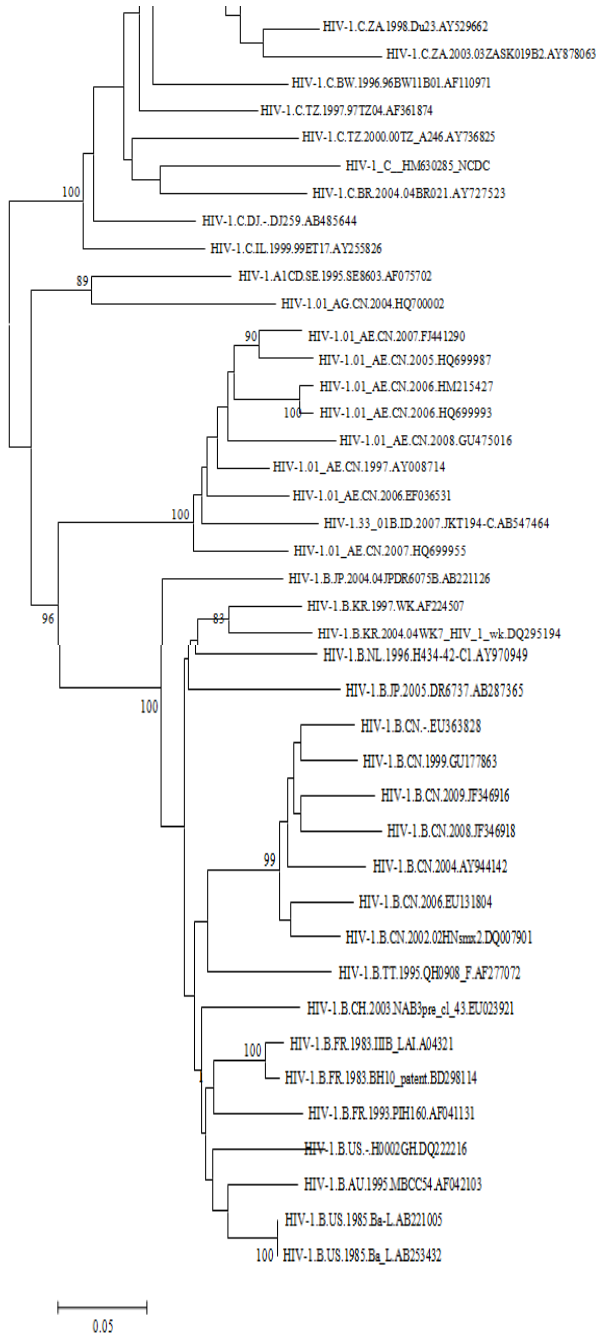
Subtype C Signatures	R	Q	T	F	-	V	N
Frequency in Subtype C	0.966	1.000	0.690	0.690	0.586	0.586	0.690
Frequency in Subtype B	0.000	0.000	0.000	0.000	0.255	0.255	0.000
Subtype B Signatures	H	R	A	Y	G	I	D
Frequency in Subtype C	0.000	0.000	0.276	0.276	0.345	0.414	0.301
Frequency in Subtype B	0.647	0.588	1.000	1.000	0.725	0.745	0.784

(a)

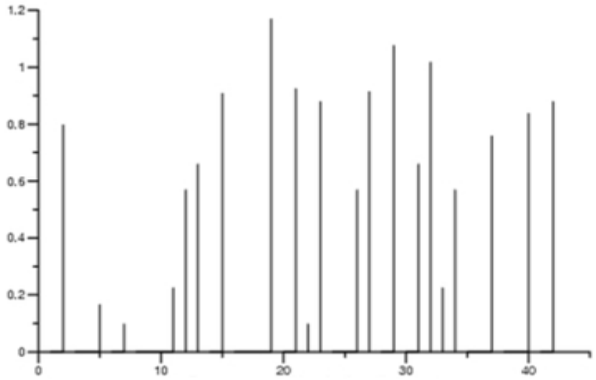
Subtype C Signatures	S	V	R	I	G	R	Q	T	A	R
Frequency in Subtype C	0.488	0.395	0.860	0.419	0.628	0.535	0.605	0.605	0.674	0.349
Frequency in Subtype B	0.111	0.056	0.419	0.194	0.028	0.000	0.000	0.083	0.306	0.000
Subtype B Signatures	R	I	H	-	-	P	R	A	T	K
Frequency in Subtype C	0.302	0.372	0.000	0.372	0.372	0.442	0.233	0.395	0.209	0.140
Frequency in Subtype B	0.583	0.833	0.361	0.778	0.778	0.750	0.944	0.500	0.611	0.444

(b)

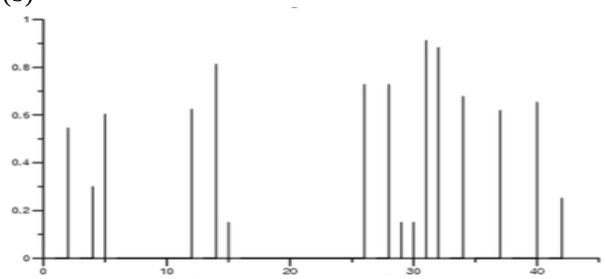



HIV-1.C.ZA.1998.Du23.AY529662
HIV-1.C.ZA.2003.03ZASK019B2.AY878063
HIV-1.C.BW.1996.96BW11B01.AF110971
HIV-1.C.TZ.1997.97TZ04.AF361874
HIV-1.C.TZ.2000.00TZ_A246.AY736825
HIV-1_C__HM630285_NCDC
HIV-1.C.BR.2004.04BR021.AY727523
HIV-1.C.DJ.-.DJ259.AB485644
HIV-1.C.IL.1999.99ET17.AY255826
HIV-1.A1CD.SE.1995.SE8603.AF075702
HIV-1.01_AG.CN.2004.HQ700002
HIV-1.01_AE.CN.2007.FJ441290
HIV-1.01_AE.CN.2005.HQ699987
HIV-1.01_AE.CN.2006.HM215427
HIV-1.01_AE.CN.2006.HQ699993
HIV-1.01_AE.CN.2008.GU475016
HIV-1.01_AE.CN.1997.AY008714
HIV-1.01_AE.CN.2006.EF036531
HIV-1.33_01B.ID.2007.JKT194-C.AB547464
HIV-1.01_AE.CN.2007.HQ699955
HIV-1.B.JP.2004.04JPDR6075B.AB221126
HIV-1.B.KR.1997.WK.AF224507
HIV-1.B.KR.2004.04WK7_HIV_1_wk.DQ295194
HIV-1.B.NL.1996.H434-42-C1.AY970949
HIV-1.B.JP.2005.DR6737.AB287365
HIV-1.B.CN.-.EU363828
HIV-1.B.CN.1999.GU177863
HIV-1.B.CN.2009.JF346916
HIV-1.B.CN.2008.JF346918
HIV-1.B.CN.2004.AY944142
HIV-1.B.CN.2006.EU131804
HIV-1.B.CN.2002.02HNsmx2.DQ007901
HIV-1.B.TT.1995.QH0908_F.AF277072
HIV-1.B.CH.2003.NAB3pre_cl_43.EU023921
HIV-1.B.FR.1983.IIIB_LAI.A04321
HIV-1.B.FR.1983.BH10_patent.BD298114
HIV-1.B.FR.1993.PIH160.AF041131
HIV-1.B.US.-.H0002GH.DQ222216
HIV-1.B.AU.1995.MBCC54.AF042103
HIV-1.B.US.1985.Ba-L.AB221005
HIV-1.B.US.1985.Ba_L.AB253432
0.05

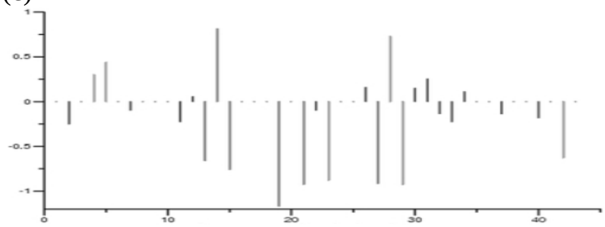
(a)



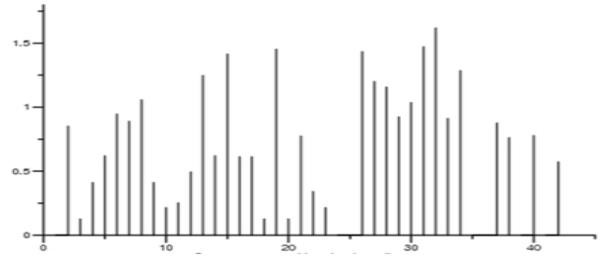
(b)



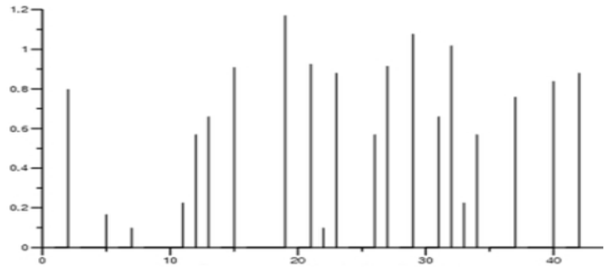
(c)



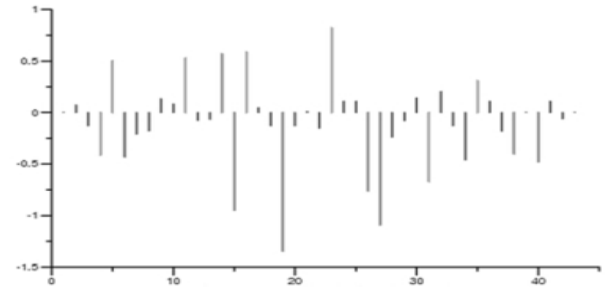
(d)



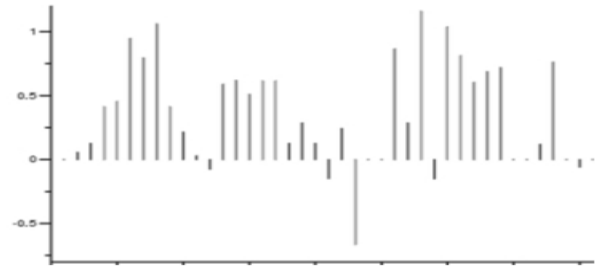
(e)



(f)



(g)



(h)

Figure 2: The above figure shows the entropy plots, constructed using 'Entropy' tool of the Los Alamos Database. On the x-axis is the amino acid position, whereas, the y-axis shows the entropy/sequence variability. The above entropy plots show sequence variability in (a) CCR5 utilizing V3 sequences of subtype C (b) CCR5 utilizing V3 sequences of subtype B (c) entropy difference between 'a' and 'b' (d) CXCR4 utilizing V3 sequences of subtype C (e) CXCR4 utilizing V3 sequences of subtype B (f) entropy difference between 'd' and 'e' (g) entropy differences between the CCR5 and CXCR4 utilizing co-receptors in subtype B sequences (h) entropy differences between the CCR5 and CXCR4 utilizing co-receptors in subtype C sequences. The sites with a significant ($p \leq 0.05$) difference in the entropy are shown in red.

Figure 3: Sequence logos of (a) CCR5 co-receptor utilizing subtype C sequences (b) CCR5 co-receptor utilizing subtype B sequences (c) CXCR4 co-receptor utilizing subtype C sequences (d) CXCR4 co-receptor utilizing subtype B sequences. Size of each character displays the proportion of sequences in which a particular amino acid occurs.

(a)

(b)

(c)

(d)