

Original Research Article

Co receptor Usage of Human Immunodeficiency Virus Type 1 Strains among Individuals Presenting for HIV Counselling and Testing in Ibadan, Nigeria

ABSTRACT

Aims To evaluate HIV-1 Coreceptor usage in Nigeria.

Study design: Cross-sectional study.

Place and Duration of Study: Department of Virology, College of Medicine, University College Hospital, Ibadan from January 2004 to December 2006.

Methodology: Genomic DNA was extracted from blood samples of 85 (42 males, 43 females; age range 18-58 years) consenting HIV-1 infected antiretroviral therapy-naïve individuals presenting at a voluntary counseling and testing centre. The *env* C2-V3 region of HIV-1 proviral DNA was amplified by nested PCR, successfully sequenced, manually edited and evolutionary relationships determined by maximum likelihood using MEGA 5.03^R from 64 of the blood samples of the HIV-1 infected patients. HIV-1 coreceptor usage was predicted based on genotypic analysis of HIV-1 *env* V3 loop sequences.

Results: Phylogenetic analysis showed HIV-1 subtypes A, G, CRF02_AG, CRF06_cpx, and CRF35_AD among the study participants. The V3 loop region of the viruses had amino acid sequence conservation in the base positions 1-8 and 26-35 and tip regions and sequence variability, including mutations and deletions at positions 9-25. Most (76.6%) of the sequences had the GPGQ crown motif while the GPGQ/L/R/K substitution was observed in 18.8%. The number of N-linked glycosylation sites ranged from 0 to 4 per *env* C2-V3 amino acid sequence with only 37.5% of the sequences having all 4 N-linked glycosylation sites. Predicted frequencies for CCR5 and CXCR4 genotypes were 31.2% and 68.8%, respectively, while 10% of the CCR5-tropic viruses showed Maraviroc-associated resistant mutations.

Conclusion: CXCR4-tropic viruses predominate among the studied population irrespective HIV-1 subtype and suggest the need for further studies involving larger sample size prior introduction of coreceptor inhibitors like Maraviroc for management of HIV infection in Nigeria.

Keywords: Coreceptor usage, HIV-1 strains, CCR5, CXCR4, C2-V3 region, V3 loop

1. INTRODUCTION

Early in infection, HIV-1 generally use the CCR5 chemokine coreceptor, along with CD4, for cellular entry [1]. In many HIV-1 infected individuals, viral genotypic changes arise that allow the virus to use CXCR4, either in addition to CCR5 or alone, as an entry receptor [2] and this change is associated with accelerated CD4 decline and more rapid progression of HIV-1 disease [3]. Based on chemokine coreceptor usage, HIV-1 viruses can be classified as CCR5-tropic, CXCR4-tropic or dual-tropic (R5X4) viruses [4]. The V3 loop of gp120 largely determines coreceptor usage and the development of CXCR4 viruses is gradual and involves the accumulation of multiple amino acid changes in the V3 [5,6].

With the introduction of HIV-1 coreceptor inhibitors as components of antiretroviral therapy [7], it is increasingly imperative to screen HIV positive individuals before they commence therapy with coreceptor inhibitors. However, there is little information regarding bioinformatics prediction of HIV-1 coreceptor usage in Nigeria, yet this information is important for drug design and clinical management of patients before coreceptor inhibitors are used for treatment. This study was therefore carried out to evaluate virus coreceptor usage by bioinformatics method based on the V3 loop sequences of HIV-1 from HIV-1 infected antiretroviral therapy-naïve individuals.

2. METHODOLOGY

2.1 Study location

This study was carried out in the Department of Virology, College of Medicine, University College Hospital, Ibadan. The department offers HIV counseling and testing services as well as monitoring of Anti-Retroviral Therapy (ART) response in HIV infected patients. The department is also the Center for the WHO National Polio laboratory, HIV reference laboratory and WHO Collaborating Center for Influenza surveillance. Patients enrolled for the study were mostly from Ibadan and other locations in the southwestern region of Nigeria.

2.2 Study population

The study population consisted of 85 consenting HIV-1 infected anti-retroviral therapy-naïve individuals. They included 42 males and 43 females with median age of 37 years (range 18-58 years) who presented at the HIV counseling and testing centre in the Department of Virology, College of Medicine, University College Hospital, Ibadan from January 2004 to December 2006. The University of Ibadan/UCH ethical review board approved the study protocol and written informed consent was obtained from every individual whose blood sample was used for the study.

2.3 Sample collection, preparation and storage

Blood sample was collected by venepuncture from each patient into sterile vacutainer blood collection tubes containing EDTA (ethylene diamine tetra acetic acid) as coagulant. Each specimen was labelled with the date of collection and laboratory identification number. Plasma was separated from each sample and both plasma and packed cells were stored at -20°C until analysed. Only samples that were positive by Western blot for HIV-1 antibodies were analysed further for the study.

2.4 Laboratory analysis

2.4.1 Serology

Initial HIV screening was by Genscreen Ultra HIV Ag-Ab (BIORAD, France) and confirmation was by the Western blot assay using New Lav Blot 1 (BIORAD, France). The results were interpreted according to the manufacturer's instructions.

2.4.2 Molecular analysis

Genomic DNA was extracted from whole blood using the QiaAmp DNA Blood Mini kit (Qiagen, Maryland, USA) according to manufacturer's instruction. The *env* C2-V3 region of HIV-1 from the genomic DNA extract was amplified by nested PCR method using primers WT1 and WT2, and KK40 and KK30 for the first and second rounds, respectively [8]. Amplified DNA was directly sequenced with the second round primers in the Genetic Analyzer 3130xl (Applied Biosystems, California, USA) in both forward and reverse directions using Sequencing kit v3.1 (Applied Biosystems, California, USA). The sequences were manually edited using the sequencing software v3.1 (Applied Biosystem, California, USA) and MEGA 5.03, and analysed for phylogenetic relationships with reference sequences from the HIV database (Los Alamos, New Mexico, USA) by ClustalW and

Maximum Likelihood methods in MEGA 5.03 software based on their nucleotide (approximately 350bp) and translated amino acid sequences [9].

2.4.3 Virus phenotype determination

The virus coreceptor usage was predicted by combined criteria using a genotypic rule based on one of the following criteria for predicting CXCR4 (R4) coreceptor usage (i) Arginine (R) or Lysine (K) at position +11 of V3 and/K at position +25, (ii) R at position +25 of V3 and a net charge $\geq +5$, or (iii) a net charge of $\geq +6$ or (iv) bioinformatic analysis by comparison between Cons B and query V3 loop sequence in **Geno2Phen** which uses the presence of X4 specific mutations in the V3 loop as a predictor of X4 phenotype [10-14].

3. RESULTS AND DISCUSSION

3.1 HIV-1 subtype

The target region of approximately 350bp was successfully sequenced from 64 (75.3%) of the 85 HIV-1 infected antiretroviral therapy-naïve individuals. The 64 HIV-1 infected individuals comprised of 34 males and 30 females with a median age of 37 years (range 18-58 years). Overall, 3.1%, 53.1%, 28.1%, 14.1% and 1.6% of the HIV-1 *env* C2-V3 nucleotide sequences were identified as subtype A, G, CRF02_AG, CRF06_cpx, and CRF35_AD, respectively [Figure 1].

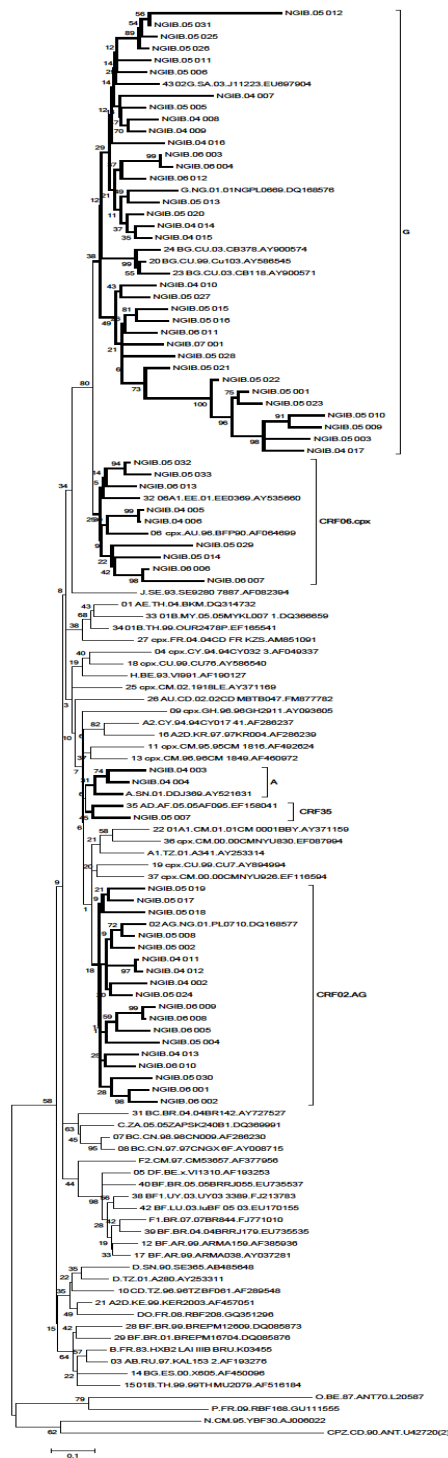


FIGURE 1: PHYLOGENETIC DISTRIBUTION OF HIV-1 STRAINS AMONG HIV-1 POSITIVE INDIVIDUALS

104 3.2 V3 loop sequence analysis

105 The sequence analysis of the V3 loop showed total conservation for 3 (P16, G17, and A33)
106 out of the 35 amino acid residues (8.6% of conservation). P16 and G17 are motifs within the
107 GPGQ motif of the V3 tip and A33 is in the base. The positions in which variation did not
108 exceed 10% were concentrated in the base (R3-4.7%, P4-7.8%, I30-4.7%, R31-7.8%), the
109 tip (G15-4.7%) and the stem (N6-3.1%, I26-6.3%). The most variable position was the 25th,
110 which was the determinant for the 11/25 rule used for predicting HIV-1 coreceptor usage.
111 The variation in the Cysteine residues limiting the V3 loop was 3.1% for C1 and 1.6% for
112 C35 [Figure 2].

113
114 The conserved GPGQ crown motif was the most common sequence observed at the V3 tip
115 occurring in 76.6% (49) of the translated HIV-1 V3 loop amino acid sequences. There was
116 substitution from GPGQ to other V3 crown motifs like GPGL (8), GPGR (3), GPGK (1),
117 APGQ (2) and RPGQ (1), as shown in figure 2.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	Subtype	Net Charge	Geno2Pheno (10%)	Coreceptor		
Cons B	C	T	R	P	N	N	N	T	-	R	K	S	I	H	I	G	P	G	R	A	F	Y	T	T	G	E	I	I	G	D	I	R	Q	A	H	C					
NGIB.04.001	C	T	R	P	G	N	N	T	-	R	R	-	Y	K	N	G	P	G	Q	S	F	Y	A	T	G	D	I	I	G	D	T	R	Q	A	H	C	G	3	R5	R5	
NGIB.04.002	C	T	R	P	N	N	N	T	-	R	K	C	-	T	Y	G	P	G	Q	T	I	Y	A	T	G	A	I	I	G	E	I	R	Q	A	H	C	CRF_02_AG	3	X4	X4	
NGIB.04.003	C	T	R	P	N	N	N	T	-	R	K	C	K	N	-	G	P	G	Q	A	F	Y	A	T	G	A	I	T	G	D	I	I	K	A	Y	C	A	4	R5	R5	
NGIB.04.004	C	T	R	P	N	N	N	T	-	R	K	C	N	N	-	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	T	R	K	A	H	C	A	3	X4	X4	
NGIB.04.005	C	Y	R	P	N	N	N	T	-	R	K	Y	I	-	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_06_cpx	2	R5	R5	
NGIB.04.006	C	Y	R	P	N	N	N	T	-	R	K	Y	I	-	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_06_cpx	2	R5	R5	
NGIB.04.007	Y	-	R	P	N	-	-	T	-	-	R	K	Y	K	N	G	P	G	Q	T	F	Y	A	T	-	-	-	I	G	N	I	R	Q	A	-	C	G	6	X4	X4	
NGIB.04.008	C	T	R	P	S	N	N	T	-	R	R	-	Y	K	N	G	P	G	Q	A	F	Y	A	T	G	A	V	T	G	D	I	R	N	A	H	C	G	4	R5	R5	
NGIB.04.009	C	T	R	P	N	N	N	T	-	-	R	K	Y	R	N	G	P	G	Q	A	F	Y	A	T	G	A	I	I	G	D	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.04.010	C	T	R	P	N	N	N	T	-	R	K	H	T	P	-	G	P	G	Q	A	F	F	A	T	G	D	I	I	G	D	I	R	E	A	H	C	G	1	R5	R5	
NGIB.04.011	C	S	R	P	G	N	N	T	-	-	R	Q	Y	T	Y	G	P	G	Q	A	F	Y	A	T	G	D	I	T	G	D	I	R	Q	A	H	C	CRF_02_AG	1	R5	R5	
NGIB.04.012	C	S	R	P	G	N	N	T	-	R	R	-	Y	T	Y	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.04.013	C	T	R	P	G	N	N	T	-	-	R	K	Y	T	Y	G	P	G	Q	V	F	Y	T	D	-	-	-	I	G	D	I	R	Q	A	H	C	CRF_02_AG	3	X4	X4	
NGIB.04.014	C	T	R	P	N	N	N	T	-	R	K	H	K	N	-	G	P	G	Q	M	F	Y	A	T	G	D	I	T	G	D	I	R	Q	A	H	C	G	3	R5	R5	
NGIB.04.015	C	T	R	P	N	N	N	T	-	-	R	K	Y	K	I	-	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	Y	C	G	3	R5	X4
NGIB.04.016	C	T	R	A	N	N	N	T	-	-	R	K	Y	K	N	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	K	Q	A	H	C	G	3	R5	X4	
NGIB.04.017	C	T	T	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.001	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.05.002	C	T	R	P	N	N	N	T	-	-	R	K	D	I	Y	G	P	G	Q	T	F	Y	A	T	G	D	I	V	G	D	I	R	Q	A	H	C	CRF_02_AG	2	R5	X4	
NGIB.05.003	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	T	I	I	G	D	I	K	Q	A	H	W	G	4	X4	X4	
NGIB.05.004	C	T	R	L	G	N	N	T	-	R	R	-	Y	T	Y	G	P	G	R	A	F	F	T	S	G	D	V	T	G	D	I	R	K	A	Y	C	CRF_02_AG	4	X4	X4	
NGIB.05.005	C	T	R	P	G	N	N	T	-	-	I	R	Y	K	N	G	P	G	Q	A	F	Y	A	T	G	K	I	T	G	D	I	R	K	A	H	C	G	5	X4	X4	
NGIB.05.006	C	T	R	P	N	N	N	T	-	R	R	-	Y	N	I	A	P	G	Q	A	F	Y	T	T	G	A	I	I	G	D	I	R	Q	A	Y	C	G	3	R5	R5	
NGIB.05.007	C	T	R	P	N	N	N	T	-	R	K	C	-	T	Y	G	P	G	Q	S	F	F	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_35_AD	2	R5	R5	
NGIB.05.008	C	T	R	P	N	N	N	T	-	R	K	C	-	T	Y	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	N	I	R	Q	A	H	C	CRF_02_AG	3	R5	R5	
NGIB.05.009	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.05.010	C	T	R	P	N	N	Y	T	-	K	K	D	I	-	I	G	P	G	L	A	F	P	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.011	C	T	R	P	G	N	N	T	-	-	R	K	D	I	-	I	G	P	G	Q	A	I	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	G	2	X4	X4
NGIB.05.012	W	T	R	P	S	I	N	T	-	-	R	I	Y	R	D	G	P	G	Q	A	V	S	A	T	V	S	I	L	V	N	L	-	I	A	D	C	G	1	X4	X4	
NGIB.05.013	C	I	R	P	N	N	N	T	-	R	K	H	K	N	-	G	P	G	Q	T	F	Y	T	T	G	A	I	I	G	N	I	R	Q	A	H	C	G	5	R5	R5	
NGIB.05.014	C	I	R	P	G	N	N	T	-	K	R	Y	I	-	L	A	P	G	Q	V	F	Y	A	T	-	-	-	I	G	D	I	R	Q	A	H	C	CRF_06_cpx	3	X4	X4	
NGIB.05.015	C	T	R	P	N	N	N	T	-	-	R	K	Y	T	T	G	P	G	Q	A	I	Y	A	T	G	A	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.016	C	S	R	P	N	N	N	T	-	-	R	K	Y	T	T	G	P	G	Q	A	I	Y	A	T	G	A	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.017	C	T	R	P	G	N	N	T	-	R	K	C	-	T	Y	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	K	A	H	C	CRF_02_AG	3	X4	X4	
NGIB.05.018	C	T	R	P	N	N	N	T	-	V	Q	G	-	A	C	G	P	G	Q	T	F	Y	A	K	D	D	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	0	R5	R5	
NGIB.05.019	C	I	R	P	S	N	N	T	-	-	R	K	F	S	Y	G	P	G	L	K	F	Y	T	T	G	N	I	I	G	D	I	R	K	A	H	C	CRF_02_AG	5	X4	X4	
NGIB.05.020	C	T	R	P	N	N	N	T	-	-	R	K	Y	K	N	G	P	G	Q	V	F	Y	A	T	G	E	I	I	G	D	I	R	Q	A	Y	C	G	3	X4	X4	
NGIB.05.021	C	T	R	P	N	N	N	T	-	R	K	Y	I	-	I	G	P	G	Q	A	L	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	G	2	R5	R5	
NGIB.05.022	C	T	R	P	N	N	Y	T	K	K	K	D	I	-	I	G	P	G	L	A	F	R	A	K	G	N	I	I	G	D	I	R	Q	A	H	C	G	5	X4	X4	
NGIB.05.023	C	T	G	P	G	N	N	T	-	K	K	Y	I	-	L	A	P	G	Q	T	F	F	A	T	G	N	I	I	G	D	I	R	Q	A	H	C	G	3	X4	X4	
NGIB.05.024	C	T	R	P	N	N	N	T	-	R	K	H	A	Y	-	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.05.025	C	T	G	P	H	N	N	T	-	R	R	-	Y	K	D	G	P	G	Q	A	F	Y	A	T	G	E	I	I	G	N	I	R	Q	A	H	C	G	2	R5	R5	
NGIB.05.026	C	T	R	P	N	N	N	T	-	R	R	-	Y	K	D	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	N	I	R	Q	A	H	C	G	3	R5	R5	
NGIB.05.027	C	T	R	P	N	N	N	T	-	-	R	K	Y	N	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	Y	C	G	2	R5	X4	
NGIB.05.028	C	T	R	P	G	N	N	T	-	R	K	D	-	D	I	G	P	G	R	V	F	Y	A	T	S	D	V	V	G	D	I	R	Q	A	H	C	G	1	X4	X4	
NGIB.05.029	C	T	R	P	N	N	N	T	-	R	K	Y	I	P	-	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	R	A	Y	C	CRF_06_cpx	3	R5	R5	
NGIB.05.030	C	T	R	P	D	N	N	T	-	-	R	Q	C	A	Y	G	P	G	Q	V	F	Y	A	N	K	D	I	I	G	N	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.05.031	C	T	R	P	N	N	N	T	-	R	R	-	Y	R	D	G	P	G	Q	A	F	Y	A	T	G	A	I	V	G	N	I	R	Q	A	H	C	G	4	X4	X4	
NGIB.05.032	C	T	R	P	N	N	N	T	-	R	K	Y	T	H	-	G	P	G	Q	A	F	Y	A	T	G	A	I	I	G	N	I	R	Q	A	H	C	CRF_06_cpx	4	R5	R5	
NGIB.05.033	C	T	R	P	N	N	N	T	-	R	R	Y	T	H	-	G	P	G	R	V	F	Y	A	T	G	D	I	I	G	D	I	R	R	A	H	C	CRF_06_cpx	4	X4	X4	
NGIB.06.001	C	S	R	P	S	N	N	T	-	R	R	D	-	R	Y	G	P	G	Q	M	F	Y	A	A	G	E	I	I	G	D	I	R	Q	A	H	C	CRF_02_AG	2	X4	X4	
NGIB.06.002	C	T	R	P	S	N	N	T	-	R	K	D	-	R	Y	G	P	G	Q	M	F	Y	A																		

Most potential N-linked glycosylation sites were conserved in the immediate vicinity of and within the V3 loop of the HIV-1 sequences analyzed. The NxS/Ty sequon for the potential N-glycosylation site centered at the 5'-Cys V3 loop residue was conserved in 46 (71.9%) of the HIV-1 C2-V3 amino acid sequences with 41 (89.1%) presenting with the NCT (Asparagine-Cysteine-Threonine) sequon [Figure 2]. Potential N-linked glycosylation sites located six amino acids downstream from the 5' cysteine residue in the V3 loop was conserved in 61 (95.3%) of HIV-1 infected individuals with 54 (88.5%) presenting with the NNT (Asparagine-Asparagine-Threonine) sequon.

Interestingly, one individual had no potential N-linked glycosylation site in the HIV-1 C2-V3 amino acid sequence analyzed. This sequence (NGIB.04__007) is a subtype G HIV-1 virus having the GPGQ V3 tip motif with a predicted X4 phenotype.

3.3 Prediction of Coreceptor usage

Using the 4 combined criteria, 68.8% (44) of V3 loop amino acid sequences had predicted X4 virus phenotype. Of these 44 V3 loop amino acid sequences, 7 V3 loops were classified as X4 phenotype based wholly on the 11/25 rule while the remaining 37 were predicted based on a combination of the other 3 criteria.

On further analysis of the V3 loop amino acid sequences that met each criterion, 16 of the V3 loop amino acid sequences had Lysine residue at position +11 while only one had Arginine residue at position +11. The only V3 loop sequence with the Arginine residue at position +11 also had a lysine residue at position +25. One V3 loop amino acid sequence had a net charge of +6. Bioinformatics analysis with Geno2Pheno at 10% false positive rate (FPR) predicted 38 (including some V3 loop sequences that met the other criteria) as X4 viruses. All V3 loop sequences with 33 and 31 amino acid residues, and those not limited by the Cysteine residues, were predicted as X4 virus phenotype. In addition, HIV-1 subtypes G and CRF02_AG had a higher prevalence of CXCR4 virus phenotypes compared to HIV-1 CRF06_cpx in the study population [Table 1]

TABLE 1. FREQUENCY OF CORECEPTOR USAGE AMONG HIV-1 SUBTYPES

	Subtype A	Subtype G	CRF02_AG	CRF06_cpx	TOTAL
Subtype	(%)	(%)	(%)	(%)	(%)
Total	2	34	18	9	63
X4 Phenotype	1(50.0)	25(73.5)	14(77.8)	4(44.4)	44(68.8)

DISCUSSION.

The increasing interest in HIV-1 tropism is related to the introduction of CCR5 inhibitors as components of antiretroviral therapy [7]. Although, the “gold standard” for characterization of HIV-1 tropism is a recombinant virus phenotypic entry assay, genotypic methods based on the V3 amino acid sequence analysis and characteristics have been successful with different HIV-1 subtypes using sequence-based criteria derived from HIV-1 subtype B [15-17]. Prediction of HIV-1 coreceptor usage by genotypic methods are easier to do, more cost effective and have been found to be useful as screening strategy in medical practice [18,19]. Moreover, the anti-CCR5 inhibitor, Maraviroc, can be taken with or without food, and does not require refrigeration and thus, can be very suitable for resource-limited settings like sub-Saharan Africa.

The presence of CXCR4-tropic HIV-1 isolates in HIV-1 infected individuals is associated with a rapid decline of CD4⁺ T cells, rapid disease progression, and reduced survival time after AIDS diagnosis [20]. Various studies on motif analysis of the V3 region of *env* have shown certain characteristics that are strongly associated with CXCR4 usage and they include: the presence and accumulation of CXCR4-associated mutations; insertions and deletions in V3, the presence of basic amino acids at V3 positions 11 and/or 25, the loss of a glycosylation site and an increased positive amino acid charge ($\geq +6$) [10-14].

Although the typical V3 loop amino acid sequence is 35 amino acids, V3 amino acid sequences of HIV-1 subtypes and CRFs found in this study showed that all, except three sequences, were 34-amino acids in length. Short length V3 amino acid sequences have been reported by other investigators and are frequently associated with CXCR4 usage [5-6]. Sequence conservation in V3 loop was almost entirely consistent when considering the first 8 amino acid residues proximal to the terminal Cysteine residue in the base region across the subtypes evaluated in this study. It has been reported that sequence conservation in the V3 region across different HIV-1 subtypes is consistent with a V3 interaction with an invariable cellular protein [21]. The sequence variability in the stem and turn regions is as a result of site specific mutations and consequently, most (7/10) of the X4-associated mutations detected in this study occur in the stem and turn regions (positions 9-25) of the V3 loop. This observation is consistent with findings by previous researchers and these site-specific mutations at either side of the V3 loop crown motif are important determinants of CXCR4 coreceptor usage [8,21].

Some authors have shown that the V3 loop of HIV-1 has a highly conserved crown motif and the GPGQ motif (which was in the majority in this study) in particular, is generally associated with African HIV isolates and it occurs irrespective of coreceptor usage [13,16]. However, HIV-1 subtype B has a GPGR crown motif irrespective of coreceptor usage while in HIV-1 subtype C, viruses using the CCR5 coreceptor have conserved GPGQ sequence tip and CXCR4 viruses have a substitution which commonly presents as GPGR [20]. In addition, the substitution of GPGQ to GPGL/R/K as observed in 18.8% of the V3 loop sequences of HIV-1 subtypes and CRFs in this study, predict CXCR4 tropism, which is in agreement with some studies [22,23].

Another determinant of CXCR4 coreceptor usage of HIV-1 isolates which had been reported from previous studies is the presence of basic amino acids at either positions 11 and/or 25 of the V3 loop. Sixteen (94.1%) of the 17 V3 loop sequences with basic amino acids at positions 11 had Arginine residues, while one sequence had Lysine and Arginine residues at positions 11 and 25, respectively. Fouchier and colleagues [20] reported a strong association between the syncytium-inducing (SI) capacity of HIV-1 and the presence of positively charged amino acid residues at positions 11 and 25 of HIV-1 V3 loop, indicating

that the V3 genotype may be useful for prediction of the viral phenotype . Also, studies of 3D molecular modeling of the V3 loops from primary isolates whose coreceptor usage was experimentally defined have revealed a charged "patch" on the surface of V3 that correlates with coreceptor usage. This V3 surface patch is positively charged in X4-tropic viruses and negatively charged or neutral in R5-tropic viruses, and is formed by two amino acids, at position 11 and at position 24 or 25; amino acids 11 and 24 or 11 and 25 contact each other in 3D space [24]. However, a study showed that sequence variability in V3 loop among subtype C viruses was not associated with the presence of basic amino acid substitutions which is in contrast to that observed in subtype B where sequence variability is associated with basic substitutions and is a determinant of altered coreceptor usage [25].

The V3 loop sequence of one sample had a net charge of +6 and multiple amino acid deletions, including the deletion of the N-linked glycosylation site. Some studies have shown that a high net charge of +5 and above in V3 loop with a loss of one or more N-linked glycosylation sites in the C2-V3 region of HIV-1 are characteristics of X-tropic viruses [23, 26-28].

Majority (68.8%) of the V3 loop sequences of the HIV-1 strains from this study had one or more X4-associated mutations including the 16 sequences that had only basic amino acid at V3 loop position 11 and/or 25. These X-4 associated mutations are consistent with report from previous studies. A review that compared different genotypic coreceptor usage prediction methods and analyses showed that the presence of mutations in the V3 loop is predictive and indicative of CXCR4 coreceptor usage [29]. Sing *et al.* [14] using univariate analyses for the prediction of HIV coreceptor use from clonal HIV V3 loop sequences identified 41 V3 mutations significantly associated with coreceptor usage. Marcelin and colleagues [30] reported that one reason for the failure of CCR5 inhibitors was the selection of resistance to CCR5 antagonists through amino acid changes in V3 loop.

Furthermore, other studies have shown that these X4-associated mutations in the V3 loop of *env* and the presence of basic amino acids at V3 positions 11 and 25 are strongly associated with syncytium induction and CXCR4 usage. Chesebro and colleagues in their studies on HIV infectivity and replication in different permissive cells, observed that infectivity and replication in macrophages was affected by the presence of basic amino acids in positions 11 and 25 of V3 with viruses harboring such mutations being T-cell tropic and syncytium inducing [10-12]. These observations were also reported by Fouchier and colleagues in their studies [20,31-32]. Hoffman *et al.* [33] showed that these basic amino acids in V3 is phenotype associated and results in the use of the CXCR4 coreceptor. Studies on HIV-1 V3 motif analysis by Jensen *et al.* [5] also observed the association between basic amino acids in V3 and CXCR4 coreceptor usage. Raymond *et al.* [19] Soulie *et al.* [34] and Ajoge *et al.* [33] in their various studies also reported that the presence of basic amino acids at V3 positions 11 and 25 was associated with CXCR4 coreceptor usage.

The 4 N-linked glycosylation sites (NGS) within the C2V3 region are present in different proportions in all, except one, of the sequences. Only 37.5% of the C2V3 sequences had the 4 NGS. The loss of one or more N-linked glycosylation sites has been shown to be associated with the development of X4 viruses [23], assist in more efficient use of CXCR4 and might be an important factor in the switch of CCR5 to CXCR4 viruses, and therefore, virus tropism [36,37]. Pollaskis *et al.* [26] also reported that the loss of an N-linked glycosylation site within the V3 region had a major influence on virus switching from R5 to X4 phenotype.

Overall, the predicted coreceptor usage using 4 combined criteria [14, 29] for the V3 loop amino acid sequences of the HIV-1 subtypes and CRFs in this study was 31.2% and 68.8%

for CCR5 and CXCR4, respectively, independent of CD4 and viral load status of the individuals, and, irrespective of subtype. The use of the combined criteria has been shown to increase both the sensitivity and the specificity of the prediction as reported by Raymond *et al.* [19] and Ajoge *et al.* [35]. The high proportion of CXCR4 in this study might be an indication that the HIV-1 infected individuals are in the advanced stage of infection associated with accelerated disease progression, or they may be exposed to anti-retroviral therapy, or both [27, 38-40].

The X4 phenotype predominates in both subtype G and CRF02_AG viruses but is not the predominant predicted phenotype in CRF06_cpx. The higher usage of CXCR4 by HIV-1G from this study is in agreement with the results of a previous study and further corroborates the evidence that HIV-1G may have a higher proportion of X4 variants [35], although more phenotypic assays for coreceptor usage for subtype G with corresponding genotypic testing needs to be done. However, the high usage of CXCR4 by HIV-1 CRF02_AG is contrary to reports from studies by Raymond *et al.* [17] and Ajoge *et al.* [35] that reported CRF02_AG HIV-1 strains to be predominantly CCR5-tropic. The prediction on CRF06_cpx shows that individuals infected with this subtype of HIV-1 may benefit from coreceptor antagonists like Maraviroc, although there is need for better characterization of HIV-1 CRF06_cpx tropism, and HIV-1G, in Nigeria.

In addition, 2 (10%) of the predicted CCR5-using HIV-1 from this study had the A22T mutation that confers resistance to CCR5 inhibitors like Maraviroc. The high prevalence of X4 tropic viruses and the presence of Maraviroc resistant mutations in some R5 viruses indicate that coreceptor usage ability of patients' viral population will clearly be required before clinical administration of Maraviroc. Moreover, if future phenotypic and genotypic studies confirm the predominance of Maraviroc-resistant CCR5, and CXCR4 phenotypes in Nigeria, a wide scale introduction and clinical application of Maraviroc in HIV-1 treatment and control will be limited.

4. CONCLUSION

This data show a high prevalence of HIV-1 subtype with the GPGQ crown motif in the V3 region suggesting a functional use of this motif. Moreover, X4 viruses predominate in the study population and it is associated with multiple amino acid deletions and mutations in V3 and the loss of one or more N-linked glycosylation sites with some CCR5-tropic viruses showing Maraviroc resistant mutations.

It will be important to carry out genotypic prediction tests for HIV-1 coreceptor usage in HIV patients' population encompassing the various regions of the country and in the predominant circulating HIV-1 subtypes before the introduction of CCR5 antagonists for clinical management of infected persons in Nigeria.

The sequences have been submitted in GenBank with Accession numbers **KF437580-KF437619**.

303 CONSENT

304
305 Written informed consent was obtained from every individual whose blood sample was used
306 for the study.

308 ETHICAL APPROVAL

309 The University of Ibadan/UCH ethical review board approved the study protocol and written
310 informed consent was obtained from every individual whose blood sample was used for the
311 study.

313 References

- 314
315 1. Bjorndal A, Sonnerborg A, Tscherning C, Albert J, Fenyo EM. Phenotypic
316 characteristics of human immunodeficiency virus type 1 subtype C isolates of
317 Ethiopian AIDS patients. *AIDS Res Hum Retroviruses*. May 1 1999;15(7):647-653.
318
- 319 2. Kupfer B, Kaiser R, Rockstroh JK, Matz B, Schneweis KE. Role of HIV-1 phenotype
320 in viral pathogenesis and its relation to viral load and CD4+ T-cell count. *J Med Virol*.
321 Nov 1998;56(3):259-263.
322
- 323 3. Koot M, Vos AH, Keet RP, et al. HIV-1 biological phenotype in long-term infected
324 individuals evaluated with an MT-2 cocultivation assay. *AIDS*. Jan 1992;6(1):49-54.
325
- 326 4. Berger EA, Doms RW, Fenyo EM, et al. A new classification for HIV-1. *Nature*. Jan
327 15 1998;391(6664):240.
328
- 329 5. Jensen MA, Li FS, van 't Wout AB, et al. Improved coreceptor usage prediction and
330 genotypic monitoring of R5-to-X4 transition by motif analysis of human
331 immunodeficiency virus type 1 env V3 loop sequences. *J Virol*. Dec
332 2003;77(24):13376-13388.
333
- 334 6. Jensen MA, van 't Wout AB. Predicting HIV-1 coreceptor usage with sequence
335 analysis. *AIDS Rev*. Apr-Jun 2003;5(2):104-112.
336
- 337 7. Dorr P, Westby M, Dobbs S, et al. Maraviroc (UK-427,857), a potent, orally
338 bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5
339 with broad-spectrum anti-human immunodeficiency virus type 1 activity. *Antimicrob*
340 *Agents Chemother*. Nov 2005;49(11):4721-4732.
341
- 342 8. Kanki PJ, Hamel DJ, Sankale JL, et al. Human immunodeficiency virus type 1
343 subtypes differ in disease progression. *J Infect Dis*. Jan 1999;179(1):68-73.
344
- 345 9. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics
346 Analysis (MEGA) software version 4.0. *Mol Biol Evol*. Aug 2007;24(8):1596-1599.
347
- 348 10. Chesebro B, Nishio J, Perryman S, et al. Identification of human immunodeficiency
349 virus envelope gene sequences influencing viral entry into CD4-positive HeLa cells,
350 T-leukemia cells, and macrophages. *J Virol*. Nov 1991;65(11):5782-5789.
351
- 352 11. Chesebro B, Wehrly K, Nishio J, Perryman S. Macrophage-tropic human
353 immunodeficiency virus isolates from different patients exhibit unusual V3 envelope
354 sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical
355 amino acids involved in cell tropism. *J Virol*. Nov 1992;66(11):6547-6554.

- 356 **12.** Chesebro B, Wehrly K, Nishio J, Perryman S. Mapping of independent V3 envelope
357 determinants of human immunodeficiency virus type 1 macrophage tropism and
358 syncytium formation in lymphocytes. *J Virol.* Dec 1996;70(12):9055-9059.
359
- 360 **13.** Holm-Hansen C, Grothues D, Rustad S, Rosok B, Pascu FR, Asjo B.
361 Characterization of HIV type 1 from Romanian children: lack of correlation between
362 V3 loop amino acid sequence and syncytium formation in MT-2 cells. *AIDS Res*
363 *Hum Retroviruses.* May 1995;11(5):597-603.
364
- 365 **14.** Sing T, Low AJ, Beerenwinkel N, et al. Predicting HIV coreceptor usage on the basis
366 of genetic and clinical covariates. *Antivir Ther.* 2007;12(7):1097-1106.
367
- 368 **15.** Jensen MA, Coetzer M, van 't Wout AB, Morris L, Mullins JI. A reliable phenotype
369 predictor for human immunodeficiency virus type 1 subtype C based on envelope V3
370 sequences. *J Virol.* May 2006;80(10):4698-4704.
371
- 372 **16.** Monno L, Brindicci G, Saracino A, et al. V3 sequences and paired HIV isolates from
373 52 non-subtype B HIV type 1-infected patients. *AIDS Res Hum Retroviruses.* Mar
374 2010;26(3):365-372.
375
- 376 **17.** Raymond S, Delobel P, Mavigner M, et al. Genotypic prediction of human
377 immunodeficiency virus type 1 CRF02_AG tropism. *J Clin Microbiol.* Jul
378 2009;47(7):2292-2294.
379
- 380 **18.** Chueca N, Garrido C, Alvarez M, et al. Improvement in the determination of HIV-1
381 tropism using the V3 gene sequence and a combination of bioinformatic tools. *J Med*
382 *Virol.* May 2009;81(5):763-767.
383
- 384 **19.** Raymond S, Delobel P, Mavigner M, et al. Correlation between genotypic
385 predictions based on V3 sequences and phenotypic determination of HIV-1 tropism.
386 *AIDS.* Sep 12 2008;22(14):F11-16.
387
- 388 **20.** Fouchier RA, Brouwer M, Broersen SM, Schuitemaker H. Simple determination of
389 human immunodeficiency virus type 1 syncytium-inducing V3 genotype by PCR. *J*
390 *Clin Microbiol.* Apr 1995;33(4):906-911.
391
- 392 **21.** Patel MB, Hoffman NG, Swanstrom R. Subtype-specific conformational differences
393 within the V3 region of subtype B and subtype C human immunodeficiency virus
394 type 1 Env proteins. *J Virol.* Jan 2008;82(2):903-916.
395
- 396 **22.** Coetzer M, Cilliers T, Ping LH, Swanstrom R, Morris L. Genetic characteristics of the
397 V3 region associated with CXCR4 usage in HIV-1 subtype C isolates. *Virology.* Dec
398 5-20 2006;356(1-2):95-105.
399
- 400 **23.** Duri K, Soko W, Gumbo F, et al. Genotypic analysis of human immunodeficiency
401 virus type 1 env V3 loop sequences: bioinformatics prediction of coreceptor usage
402 among 28 infected mother-infant pairs in a drug-naive population. *AIDS Res Hum*
403 *Retroviruses.* 2011 2011;27(4):411-419.
404
- 405 **24.** Cardozo T, Kimura T, Philpott S, Weiser B, Burger H, Zolla-Pazner S. Structural
406 basis for coreceptor selectivity by the HIV type 1 V3 loop. *AIDS Res Hum*
407 *Retroviruses.* Mar 2007;23(3):415-426.
408

- 409 **25.** Ping LH, Nelson JA, Hoffman IF, et al. Characterization of V3 sequence
410 heterogeneity in subtype C human immunodeficiency virus type 1 isolates from
411 Malawi: underrepresentation of X4 variants. *J Virol.* Aug 1999;73(8):6271-6281.
412
- 413 **26.** Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA. N-linked
414 glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant
415 of CCR5 and CXCR4 coreceptor utilization. *J Biol Chem.* Apr 20
416 2001;276(16):13433-13441.
- 417 **27.** Johnston ER, Zijenah LS, Mutetwa S, Kantor R, Kittinunvorakoon C, Katzenstein
418 DA. High frequency of syncytium-inducing and CXCR4-tropic viruses among human
419 immunodeficiency virus type 1 subtype C-infected patients receiving antiretroviral
420 treatment. *J Virol.* Jul 2003;77(13):7682-7688.
421
- 422 **28.** Kaleebu P, Nankya IL, Yirrell DL, et al. Relation between chemokine receptor use,
423 disease stage, and HIV-1 subtypes A and D: results from a rural Ugandan cohort. *J*
424 *Acquir Immune Defic Syndr.* May 1 2007;45(1):28-33.
425
- 426 **29.** Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of
427 HIV coreceptor usage. *Nat Biotechnol.* Dec 2007;25(12):1407-1410.
428
- 429 **30.** Marcelin AG, Ceccherini-Silberstein F, Perno CF, Calvez V. Resistance to novel
430 drug classes. *Curr Opin HIV AIDS.* Nov 2009;4(6):531-537.
431
- 432 **31.** Fouchier RA, Brouwer M, Kootstra NA, Huisman HG, Schuitemaker H. HIV-1
433 macrophage tropism is determined at multiple levels of the viral replication cycle. *J*
434 *Clin Invest.* Nov 1994;94(5):1806-1814.
435
- 436 **32.** Fouchier RA, Groenink M, Kootstra NA, et al. Phenotype-associated sequence
437 variation in the third variable domain of the human immunodeficiency virus type 1
438 gp120 molecule. *J Virol.* May 1992;66(5):3183-3187.
439
- 440 **33.** Hoffman NG, Seillier-Moisewitsch F, Ahn J, Walker JM, Swanstrom R. Variability in
441 the human immunodeficiency virus type 1 gp120 Env protein linked to phenotype-
442 associated changes in the V3 loop. *J Virol.* Apr 2002;76(8):3852-3864.
443
- 444 **34.** Soulie C, Malet I, Lambert-Niclot S, et al. Primary genotypic resistance of HIV-1 to
445 CCR5 antagonists in CCR5 antagonist treatment-naïve patients. *AIDS.* Oct 18
446 2008;22(16):2212-2214.
447
- 448 **35.** Ajoge HO, Gordon ML, de Oliveira T, et al. Genetic characteristics, coreceptor
449 usage potential and evolution of Nigerian HIV-1 subtype G and CRF02_AG isolates.
450 *PLoS One.* 2011;6(3):e17865.
451
- 452 **36.** Nabatov AA, Pollakis G, Linnemann T, Kliphuis A, Chalaby MI, Paxton WA.
453 Inpatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and
454 V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC
455 chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies. *J Virol.*
456 Jan 2004;78(1):524-530.
457
- 458 **37.** Polzer S, Dittmar MT, Schmitz H, Schreiber M. The N-linked glycan g15 within the
459 V3 loop of the HIV-1 external glycoprotein gp120 affects coreceptor usage, cellular
460 tropism, and neutralization. *Virology.* Dec 5 2002;304(1):70-80.
461

- 462 **38.** Westby M, van der Ryst E. CCR5 antagonists: host-targeted antivirals for the
463 treatment of HIV infection. *Antivir Chem Chemother.* 2005;16(6):339-354.
464
- 465 **39.** Westby M. Resistance to CCR5 antagonists. *Curr Opin HIV AIDS.* Mar
466 2007;2(2):137-144.
467
- 468 **40.** Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in coreceptor
469 use correlates with disease progression in HIV-1--infected individuals. *J Exp Med.*
470 Feb 17 1997;185(4):621-628.
471
472

473

474