Emergence of HIV-1 C/A1 and C/A1/D circulating recombinant forms, and dominance of subtype C and R5 use from whole genome sequence analysis in Addis Ababa

Melaku Adal (melakuadal@gmail.com)
Addis Ababa University College of Natural Sciences https://orcid.org/0000-0001-5130-4678
Kate El Bouzidi
University College London
Adane Mihret
Armauer Hansen Research Institute
Rawleigh Howe
Armauer Hansen Research Institute
Abraham Aseffa
Armauer Hansen Research Institute
Melanie J. Newport
Brighton and Sussex Medical School
Jaime H. Vera
Brighton and Sussex Medical School
Beyene Petros
Addis Ababa University College of Natural Sciences

Research article

Keywords: HIV, tropism, subtypes, circulating recombinant forms, Addis Ababa

Posted Date: July 23rd, 2019

DOI: https://doi.org/10.21203/rs.2.11841/v1

Abstract

Background The HIV pandemic in Ethiopia is dominated by subtype C with sporadic A and D epidemiology. The presence of subtypes A and D may result in emergence of recombinant viruses, and increase the genetic diversity that makes monitoring the HIV epidemic, and the development of vaccines and therapeutics difficult. This study is aimed at determining subtypes, circulating recombinant forms (CRFs), and the dominant coreceptor use in Addis Ababa, Ethiopia. Methods Participants with a range of purposely selected CD4+ T-cell counts were included. Chi-square and Mann-Whitney tests were used. Whole genome next-generation sequencing (NGS) of HIV was performed using a PCR amplification method and Illumina MiSeq. Subtyping and scanning of recombination were done by the REGA subtyping tool version 3.0. Prediction of coreceptor usage was performed using Geno2Pheno clonal-model and PhenoSeq. Signature amino acids and positive charges were also used in the tropism prediction. Phylogenetic analyses were conducted with MEGA version 6 using maximum likelihood with the neighbor-joining (N-J) methods. Results Sixty participants were included with a median age of 34.5 [interquartile range (IQR) 30.0-40.0]. Seven (11.7%) of the study participants were at WHO clinical stage 3/4 and 13 (21.7%) were at AIDS stage with CD4+ T cell count <200 cells/uL. Among the total 60 HIV genomes sequenced, 49 were subtype C (81.7%), one was subtype A1 (1.7%), six were recombinant C/A1 (10%), three were recombinant C/A1/D (5.0%), and one was unassigned. From 50 of the sequences where coreceptor usage was determined by PhenoSeq, 44 (88.0%) were CCR5-tropic and six (12.0%) used CXCR4. Conclusion The study confirmed that the dominant subtype in Addis Ababa is HIV-1 subtype C. In addition, HIV-1 subtype A1, CRFs C/A1 and C/A1/D were also identified. The dominance of R5-tropic viruses was detected and these were associated with a higher CD4 T-cell count and lower viral load. Further studies on HIV subtypes and CRFs will be essential to fully understand HIV/AIDS epidemiology. In addition, the tropism information is important in Ethiopia if the use of the co-receptor antagonist maraviroc is planned.

Introduction

Current HIV prevalence data for Ethiopia [1], estimate there are 610,335 people living with HIV (PLHIV) with the adult HIV prevalence being around 1%. The HIV epidemic in Ethiopia is dominated by subtype C [2, 3, 4], which also accounts for more than 50% of the global pandemic [5]. However, sporadic infections with other subtypes, A and D have also been reported [2, 3]. Subtypes A and D predominate in East and central African countries such as Kenya, Uganda, and Tanzania [6, 7, 8]. In Djibouti subtypes A and C are reported [9]. Although effort was made to detect the existence of recombinants by undertaking the first full length subtype C sequence from a 1986 Ethiopian sample [10], the first evidence of subtype A/C recombinant from a 1991 Addis Ababa sample was reported in 1998 by Sherefa *et al.* [11].

Genetic subtypes may differ in important biological properties such as virulence, tissue tropism and transmissibility [12]. Ethiopian patients with HIV–1 subtype C harbor a remarkably low frequency of syncytium inducing (SI) CXCR4 phenotype viruses [13]. There is a strong correlation between the viral tropism and progression to disease [14, 15]. Non-syncytium inducing (NSI) HIV–1 strains use primarily

CCR5 computing with -chemokines including regulated on activation, normal T expressed and secreted (RANTES), macrophage inflammatory protein–1α or 1β (MIP–1 and MIP–1) receptor, while SI strains use CXCR4 in comptetition with -chemokines, for example, stromal differentiating factor 1 (SDF–1) [16]. For prediction of coreceptor usage, different bioinformatics algorithms are developed [17, 18] and combining the presence of lysine and argenine amino acids at positions 11/24/25 and the net charge of V3 tested for HIV–1 is also used [19].

Those individuals who are homozygous for CCR5 delta32 deletion are relatively resistant to HIV infection, which makes CCR5 one of the therapeutic targets. However, those individuals who are heterozygous for the deletion have reduced expression of CCR5 and have slower declines in the CD4 T cell count and slower progression to AIDS [20, 21, 22]. The CCR5 antagonist Maraviroc has been previously used for salvage therapy in those who have failed first and second-line treatment regimens, however it has also been evaluated as an antinflammatory agent with potential use in liver steatosis and cognitive impairment with positive results [23, 24]. Therefore, generating information on the type of the viral strain dominant in people living with HIV is important to determine the feasibility of CCR5 antagonist use in the Ethiopian context.

The cause of HIV diversity is mainly accumulation of point mutations introduced by the error prone HIV–1 reverse transcriptase during replication [25, 26]. This is amplified by the high rate of replication the virus has, where about 10^{10} virions are produced each day thereby increasing the rate of error introduction [27]. The most strongly conserved residues in the V3 loop are the two-cysteine residues, GPGX motif at the tip of the V3, and the n-linked glycosylation site adjacent to the first cysteine residue. The number of charges and glycosylation in the V3-loop can affect cellular and neutralization abilities of antibodies [28]. The extensive genetic diversity of the variants within an individual overtime and the emergence of recombinant viruses have made the development of medical interventions much more difficult. This may also enable HIV both to overcome the immune response and to develop resistance to antiviral agents. In addition, it makes difficult development of vaccine(s), diagnostics and therapeutics [29, 30, 31]. An indepth study of HIV genetic variation and classification of subtypes together with better understanding of its circulating recombinant forms (CRFs) and other recombinant genomes would be necessary for monitoring the HIV/AIDS epidemics and understanding its epidemiology. Therefore, this study was aimed at determining the HIV–1 subtypes and the dominant coreceptor tropism of the viruses circulating in Addis Ababa.

Methods

Study setting, design and population

A whole genome sequencing of HIV–1, on selected 60 samples taken from drug naïve 594 cross sectional study participants [32, 33], was done to determine subtypes, CRFs and the tropism and glycosylation site diversity. The samples were collected from four hospitals participating in the study, Addis Ababa, Ethiopia. These included the All African Leprosy Rehabilitation and Training Centre (ALERT),

and Saint Paul, Yekatit–12 and Zewditu hospitals. Samples were selected based on their CD4 T cell count and WHO stage categories <200 and 200–349 with WHO stages 1–4, and 350–500 and >500 cells/uL with WHO stages 1–2. Study participants were selected from a treatment-naïve HIV-positive cohort in Addis Ababa, Ethiopia. Participants were included if they had been followed up in clinic for at least three years and had a stored plasma sample with HIV–1 RNA >1,000 copies/mL available for sequencing. We were able to get those study participants ART naïve because test and treat was not started yet. Ethical approval for the study was obtained from Institutional Research Ethics Review Committees of the participating institutions and the Ministry of Science and Technology, Ethiopian with renewed ethical approval reference number 3.10/004/2015 (S1 Figure 1). Written informed consent was obtained from all study participants.

Haematological and virological assays: Whole blood was collected in ethylenediaminetetraacetic acid (EDTA) vacutainer tubes and plasma separated and stored at –80˚C at Armauer Hansen Research Institute (AHRI). CD4+ T-cell count was enumerated using FACSCount (Becton and Dickinson, San Jose, CA, USA) from whole blood. Abbott HIV–1 assay (Abbott Molecular Inc., Des Plaines, IL, USA) was used to determine HIV ribonucleic acid (RNA) load in 200 μL plasma.

Socio-demographic and clinical data analysis: Baseline socio-demographic and laboratory characteristics (gender, age, WHO clinical stage, CD4+ T-cell count, HIV RNA load) were analysed. Descriptive analyses included frequencies of categorical variables, median and interquartile range (IQR) for continuous variables were carried out. Chi-square test was used to test the association between categorical variables. Mann-Whitney test assessed differences of continuous variables between two categories.

HIV–1 RNA amplification and sequencing: RNA was extracted and purified from the plasma samples with HIV RNA load. The extracted RNA was reverse transcribed and the complementary deoxynucleic acid (cDNA) was processed by the polymerase chain reaction-amplicon (PCR-amplicon) method to create a template library. Next generation sequencing (NGS) was performed using an Illumina MiSeq. The sequences were generated using a PCR method with four overlapping amplicons spanning the whole genome [34].

Genome assembly: Quality control checks and trimming were done on the raw reads, followed by assembly of consensus sequences and mapping of reads onto the consensus for variant calling. Reads that were low quality and below a minimum length (50 bases) were trimmed and aligned to both human and HIV genomes. Those HIV raw reads that aligned with the human genome were discarded to avoid contamination. After this preparation, NGS raw reads were assembled into genomes using the iterative viral assembler (IVA) method

[35]. The contigs produced using IVA were aligned to an HIV sequence database which contains full length HIV genomes in order to select reference to fill gaps. The contigs were mapped and aligned to this selected reference genome and the draft genomes were constructed. Then, gap filling was done by aligning the good quality reads onto the draft genome, and replacement of bases from the reference with

those from the reads. Finally, gap filling was repeated for a maximum of 10 times to get the final consensus genome [35].

Assembled sequence analysis: Consensus sequence analysis was done to identify subtypes, CRFs, tropism and resistance mutations. Subtyping and scanning of recombination was done by the REGA subtyping tool V3.0 [36]. The V3 loop sequence was derived by gene cutter program from the whole genome [37]. Prediction of coreceptor usage was performed using Geno2Pheno clonal–model [17] and PhenoSeq [18]. In the 35 amino acids V3 loop, predicted positive charges <5.0 suggest that the virus is macrophage tropic, whereas positive charges 5.0 suggest that the virus is T cell tropic [19]. Lysine and arginine amino acids at positions 11, 24, and 25 in the V3 loop are defined and used as signature amino acids for the determination of SI and NSI phenotypes [38]. By using Geno2Pheno, a false positive rate (FPR) below 10% was considered as X4-tropic strains. PhenoSeq was used to identify the virus as X4 using or non-X4 using. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 6 [39] with inclusion of reference sequences from A, C, and D subtypes and recombinants of these subtypes. The tree was generated using maximum likelihood with the neighbor-joining (N-J) methods with 1000 bootstrap replications. Stanford HIV Drug Resistance Database was used to detect drug resistance mutations [40].

Results

Study population characteristics

A total of 60 participants, 19 (31.7%) men and 41 (68.3%) women, were selected. In addition, 13 (21.7%) samples had <200 CD4+ T cell counts and fell under the World Health Organization (WHO) clinical stages 1-4; 14 (23.3%) samples had 200-349 CD4+ T cell count and WHO clinical stages 1-4; 12 (20.0%) samples had 350-500 CD4+ T cell counts and WHO clinical stages 1-2; and 21 (35.0%) samples had >500 CD4+ T cell count and WHO clinical stages 1-2. The median age of the study participants was 34.5 (IQR: 30.0-40). Seven (11.7%) of the study participants were at WHO clinical stage ¾ and 13 (21.7%) were at AIDS stage with CD4+ T cell count <200 cells/uL (Supplement 1). WHO clinical stages were not associated with HIV-1 tropism ($p > 0.05$). However, the four groups of CD4+ T cell counts (<200-, 200-349-, 350-500- and >500 cells/uL) and HIV RNA load (<10,000- and ≥10,000 copies/mL) were associated with viral tropism ($p < 0.05$). The median viral load in all 60 study participants was 8704.0 copies/mL (IQR: 3280.75-106955.50 copies/mL). And, the median CD4+ T cell count was 368.5 cells/uL (IQR: 251.25 - 571.75 cells/uL). The median values of CD4+ T cell count and HIV RNA load between phenotypes determined by Geno2Pheno were compared using Mann-Whitney test. It is found that

[CXCR4 median 209.00 cells/uL (IQR: 126.00-288.50 cells/uL); R5 median 364.00 cells/uL (IQR: 261.50-525.00 cells/uL); p = 0.003] for CD4+ T cell count, and [CXCR4 median 107267.00 copies/mL (IQR: 74976.50-475570.00 copies/mL); R5 median 8734.00 copies/mL (IQR: 5092.50-106646.00 copies/mL); p = 0.003] for HIV RNA load (Table 1).

Table 1. Summary of demographic, hematological and virological characteristics of study participants.

Variables	Number (%)
Gender	
Male	19 (31.7)
Female	41 (68.3)
Age (years)	
18-29	9 (15.0)
30-39	34 (56.7)
40-79	17 (28.3)
WHO clinical stage	
Stage 1	27 (45.0)
Stage 2	26 (43.3)
Stage 3/4	7 (11.7)
CD4+ cell count (cells/mm^3)	
<200	13 (21.7)
≥200	47 (78.3)
HIV RNA load (copies/mL)	
<10000	33 (55.0)
≥10000	27 (45.0)

Subtyping and circulating recombinant forms

Among the total 60 HIV whole genomes sequenced, 49 were subtype C (81.7%), one was subtype A1 (1.7%) six were recombinant C/A1 (10%) and three were recombinant C/A1/D (5.0%) were identified. One of the sequences could not be assigned to a subtype by REGA (Figure 1 & 2; Table 2). The phylogenetic tree generated by the N-J method indicated that all viruses of 58 sequences belonged to subtype C, except for the virus from subject AL-062 who was infected with HIV-1 subtype A. The branch length in the phylogenetic tree between sequences of plasma was not the same in all subjects as indicated in the tree (Figure 2).

Figure 1. Percentage of subtypes, circulating recombinant forms, and R5 and CXCR4 coreceptor uses.

Tropism and glycosylation sites in V3 loop

Among the total 60 HIV genomes sequenced, 10 (16.7%) of them were not with good sequence in V3 loop to determine the coreceptor usage (tropism). From 50 of the sequences where coreceptor usage was determined by PhenoSeq, 44 (88.0%) and 6 (12.0%) were R5 and X4, respectively. In addition, phenotype by Geno2Pheno showed that 41 (82.0%) and 9 (18.0%) were R5 and X4, respectively. There was no any positively charged amino acid arginine (R) or lysine (K) at position 11. At position 24, lysine (K) was recorded for ZM-028 with net positive charge of +5. At position 25, lysine was recorded for AL-108 with the net positive charge of +5 and SP-065 with net positive charge of +6. Therefore, all three of the samples (AL-108, SP-065 and ZM-028) carried X4 tropic viruses. The other possible X4 phenotype virus with net positive charge of +5 was identified in SP-064 (Figure 1, Table 2). The NNNT 35 (70%) motif at the beginning of the loop was identified as the most dominant potential N-linked glycosylation site. Furthermore, it was found that GNNT 8 (16%), SNNT 4 (8%), GNNI 1 (2%), NNNR 1 (2%) and QNNT 1 (2%) were also identified as additional glycosylation sites (Table 2).

Table 2. HIV-1 V3 loop consensus sequences, co-receptor use, and subtypes and circulating recombinant forms.

Env name	V3 HIV aligned sequences	charge	Coreceptor usage		Subtype/recombinant
			PhenoSeq	Geno2Pheno	
	* 11 24 25 *				
AL-008	CTRPSNNTRK S IRIGPGQAFYAT G D VTGDIRQAHC	+3	R5	R5	Subtype C
AL-017	CTRPNNNTRE S IRIGPGQTFYAT G A IIGDIRQAHC	+2	R5	R5	Subtype C
AL-045	CTRPGNNTRE S IRIGPGQAFYAT G D IIGDIRQAYC	+1	R5	R5	Subtype C
AL-062	CTRPSNNTRT S IRIGPGQAFFAT G D IIGDIRQAHC	+2	R5	R5	Subtype A (A1)
AL-075	CIRPNNNTRK S MRIGPGQTFYAT G G IIGDIRAAYC	+4	R5	R5	Subtype C
AL-077	CTRPNNNTRK S VRIGPGQTFYAT G D IIGDIRQAYC	+3	R5	R5	Subtype C
AL-105	CTRPNNNTRK S IRIGPGQTFYAT G E IIGDIREAHC	+2	R5	R5	Subtype C
AL-108	CTRPNNNTIK S MRIGPGQTFYAT G K IVGNIRQAHC	+5	CXCR4	CXCR4	Subtype C
AL-115	CTRPNNNTRR S IRIGPGQAFYAT E D VIGDIRQAYC	+2	R5	CXCR4	Subtype C
AL-124	CIRPGNNTRK S IRIGPGQTFYAT G D IIGNPRKAYC	+4	R5	R5	Subtype C
AL-127	CTRPNNNTRQ S IRIGPGQTFYAT G E IIGDIRQAHC	+2	R5	R5	Subtype C
AL-128	CTRPNNNTRK S MRIGPGQTFYAT - D IIGDIRQAHC	+3	R5	R5	Subtype C
AL-134	CVRPNNNTRK S IRIGPGQAFYAT G A IIGDIRQAYC	+4	R5	R5	Subtype C
Al-136	CTRPNNNTRK S VRIGPGQAFFAT G D IIGDIRQAHC	+3	R5	R5	Subtype C
AL-137	CMRPGNNTRT S VRIGPGQTFYAT G D IVGDIKQAHC	+2	R5	R5	Subtype C
AL-149	CARPNNNTRK S VSVGPGQAIYAT G D IIGDIRQAHC	+2	CXCR4	R5	Subtype C
AL-161	CTRPNNNTRK S VRIGPGQTFYAT G A IIGEIRQAHC	+4	R5	R5	Subtype C
AL-182	CTRPNNNTRK S IRIGPGQTFYAT - D IIGDIRQAHC	+3	R5	R5	Subtype C
AL-185	CTRPNNNTRQ S IRIGPGQTFYAT G E IIGDIRQAHC	+2	R5	R5	Subtype C
AL-205	CTRPNNNTRE S IRIGPGQTFYAT G D IIGDIRQAHC	+1	R5	R5	Subtype C
SP-008	CIRPNNNRRK S VRIGPGQTFYAT G D IIGDIRAAFC	+4	R5	R5	Subtype C
SP-013	CTRPGNNTRK S VRIGPGQTFYAT G D IIGDIKQAHC	+3	R5	R5	Subtype C
SP-022	CTRPSNNTRK S VRIGPGQTFYAT G D IIGNIRQAYC	+4	R5	R5	Subtype C
SP-052	CTRPNNNTRE S IRIGPGQTFYAT G D IIGDIRQAYC	+1	R5	R5	Subtype C
SP-064	CTRPNNNTRK Y VRIGRGQVFHAT G E IIGDIRKAYC	+5	CXCR4	CXCR4	Subtype C
SP-065	CTRPNNNTRK S VRIGPGQTFYTT - K IIGNIRLAHC	+6	CXCR4	R5	Recombinant of C, A1
SP-067	CTRPGNNTRE S VRIGPGQAFYAT G E IIGDIRKAHC	+2	R5	R5	Subtype C
SP-073	CTRPNNNTRK S VRIGPGQTFFAT G E IIGNIRKAYC	+5	R5	R5	Subtype C
SP-078	CTRPGNNTRK S VRIGPGQTFYAT G D IIGDIRQAHC	+3	R5	R5	Not identified
SP-079	CTRPNNNTRK S VRIGPGQTFYAT G A IIGEIRQAHC	+5	R5	R5	Subtype C
SP-090	CTRPNNNTRK S VRIGPGQVFYAT G D IIGDIRQAHC	+3	R5	R5	Recombinant of C, A1, D
SP-095	CTRPNNNTRR S VRIGPGQTFYAT G E IIGDIKQAHC	+3	R5	R5	Subtype C
SP-109	CTRPGNNIRK S MRIGPGQAFYAT G D IIGDLRQAHC	+3	R5	CXCR4	Subtype C
SP-146	CTRPNNNTRQ S MRIGPGQAFYAM G D IIGDIRQAHC	+2	R5	R5	Subtype C
SP-151	CTRFNNNTRK S IRIGPGQAFYTA G E IIGEIRQAHC	+3	R5	R5	Subtype C
YK-003	CTRPGNNTRK S VRIGPGQTFYAT G A ----------	-	R5	CXCR4	Subtype C
YK-019	CTRPGNNTRR S VRIGPGQTFYAT G D IIGDIRQAHC	+3	R5	R5	Subtype C
YK-052	CTRPQNNTRR S VRIGPGQAFYTT G D IIGDIRQAHC	+3	R5	R5	Subtype C
ZM-005	CTRPNNNTRK S IRIGPGQAFYAR G D IIGDIRQAHC	+4	R5	CXCR4	Subtype C
ZM-019	CTRPNNNTRK S MRIGPGQVFYAT E D IIGDIRQAHC	+2	R5	CXCR4	Subtype C
ZM-028	CMRPNNNTRK S IRIGPGQTFYAT K D IIGNIRQAHC	+5	CXCR4	CXCR4	Subtype C
ZM-058	CTRPNNNTRE S VRIGPGQTFFAT G D IIGDIRQAHC	0	R5	R5	Subtype C
ZM-075	CTRPGNNTRR S VRIGPGQTFFAT G E IIGDIRQAYC	+3	R5	R5	Subtype C
ZM-086	CTRPNNNTRK S VRIGPGQTFYAT G D IIGNIRQAHC	+4	R5	R5	Subtype C
ZM-102	CTRPNNNTRT S IRIGPGQSFHAT G A ITGRIRQAHC	+4	CXCR4	R5	Subtype C
ZM-121	CTRPNNNTRK S VRIGPGQAFYAT G D IIGNIRQAYC	+4	R5	R5	Subtype C
ZM-134	CTRPNNNTRK S VRIGPGQTFYAT G D IIGNIRQAHC	+4	R5	R5	Subtype C
ZM-136	CERPNNNTRE S IRIGPGKTFYAT G E IIGDIRQAYC	+1	R5	R5	Subtype C
ZM-151	CTRHSNNTRK S IRIGPGQAFFAT G E VIGDIRLAHC	+3	R5	R5	Subtype C
ZM-156	CMRPNNNTRK S IRIGPGQAFFAT G A VTGDIRQAHC	+4	R5	CXCR4	Subtype C

Detection of drug resistance-associated mutations

No drug resistance-associated mutations were detected by the Stanford HIV Drug Resistance Database.

Figure 2. Neighbor-joining tree demonstrating the evolutionary relationship and the distance of the HIV-1 genome consensus sequences. Sixty sequences from plasma samples, subtype C, Subtupe A1, Subtype D and AC, AD, CD, ACD circulating recombinant forms as reference sequences from the Los Alamos database were used. The scale bar represents a genetic distance of 2%.

Discussion

In this study, subtype C was the dominant subtype identified. In addition, subtype A1, circulating recombinant forms C/A1 and C/A1/D were also reported. This finding is in concordance with other studies that indicated subtype C is the dominant HIV–1 variant circulating in Ethiopia [4, 41, 42]. But, sporadic infections with other subtypes, A and D have also been reported [2, 3, 19]. The reporting of CRFs is in this study is an indication of the emergence of recombinants of the predominant subtypes A, C and D circulating in East African region. The emergence of such recombinants is plausible considering the frequently reported subtypes in Ethiopia and the possibility of influx of other subtypes from the neighboring countries [6, 7, 8, 9].

This study also showed that there is a dominant use of CCR5 coreceptor. This finding is also in concordance with other studies that showed subtype C differs from the other subtypes by its lack of ability to use coreceptors other than CCR5 [13, 43, 44], and HIV–1 subtype C uses CCR5 coreceptor for cell entry frequently even in patients with advanced immunodeficiency [45, 46, 47]. Preferential transmissibility of certain NSI isolates compared with more pathogenic SI isolates may be one explanation for this finding. However, it is also possible that the primary immune response after HIV infection of an individual might be more efficient in eliminating SI viruses than in eliminating NSI viruses [38]. The other reason may be the differential expression of RANTES, MIP–1α and MIP–1β that compete with HIV for access to cell surface CCR5, and SDF–1 by competing with CXCR4 [16, 48]. Large differences were detected in determining the coreceptor usage of the subtype C in the bioinformatics tools used in this study. This discrepancy is likely due to the use of different statistical models in how to handle insertions, deletions and ambiguous positions [49].

In this setting, HIV–1 C subtype R5-tropic viruses predominate. Those with X4-tropic infections were more likely to have lower CD4+ cell counts and higher viral loads. Positively charged signature amino acids at

positions 11, 24, and 25 [38], the net positive charge ≥5.0 [19], and the potential N-linked glycosylation site within the V3 loop [28] are predictive markers for T cell tropism of the viral isolates. The net charge of the V3 loop and the lack of positively charged amino acids at positions 11, 24, and 25 indicated that almost all study subjects carried NSI viruses [38]. This finding will have clinical relevance under the circumstance when the CCR5-receptor antagonist maraviroc is decided for use in Ethiopia.

The amino acid changes in the charged V3 loop that determines cellular tropism and glycosylation differences that result in escaping from its recognition by neutralizing antibodies in the V3 loop can be the result of different immune pressure or differences in coreceptor usage [28, 50]. This may affect the transmission of the virus and leads to disease progression in the presence of neutralizing antibodies [14, 15]. In addition, the branch length in the phylogenetic tree between sequences was different that indicates the high genetic diversity within the dominant subtype C. This indicates that there is need of considering this genetic diversity in the development of vaccines. These differences could also be the factors responsible for viral escape to immunity and responsible to challenges in the development of efficacious vaccine(s). The predicted phenotypes using bioinformatics tools, signature amino acids and net positive charge confirm the low prevalence of CXCR4 usage. This is observed in HIV–1 subtype C from Ethiopian AIDS patients in some studies in contrast to other HIV–1 subtypes [13, 43, 44].

In conclusion, the epidemic in Addis Ababa is still dominated by HIV–1 subtype C. In addition, HIV–1 subtype A1, circulating recombinant forms C/A1 and C/A1/D are also identified. Therefore, continuous studies on HIV genetic variation, subtypes and CRFs will have paramount importance to understand HIV/AIDS epidemiology, vaccine design, and detection of genetic determinants related to a particular HIV. Furthermore, the dominance of R5-tropic viruses was also detected. This is important in Ethiopia if the use of the co-receptor antagonist maraviroc is planned for use in the future with a high FPR% to decrease the risk of using the CCR5-antagonist maraviroc in patients with X4 virus.

Abbreviations

AIDS = Acquire Immunodeficiency syndrome

ALERT = Leprosy Rehabilitation and Training Centre

ART = antiretroviral therapy

cDNA = Complementary deoxynucleic acid

CRFs = circulating recombinant forms

EDTA = Ethylenediaminetetraacetic acid

FPR = False positive rate

HIV–1 = Human immunodeficiency virus type 1

IQR = Interquartile range

IRERC = Institutional Research Ethics Review Committee

IRERC = Institutional Research Ethics Review Committee

IVA = Iterative viral assembler

MIP–1 and MIP–1 = macrophage inflammatory protein–1α or 1β

NGS = Next generation sequencing

N-J = Neighbor-joining

NSI = Non-syncytium inducing

PCR-amplicon = Polymerase chain reaction-amplicon

PLHIV = people living with HIV

RANTES = -chemokines regulated on activation, normal T expressed and secreted

SDF–1 = -chemokine stromal differentiating factor 1

SI = syncytium inducing

WHO = World Health Organization

Declarations

AVAILABILITY OF DATA AND MATERIALS

The consensus sequences are deposited in GenBank - submission #2238319.

COMPETING INTERESTS

Authors of this study declared that no potential conflict of interest relevant to this article.

AUTHORS' CONTRIBUTIONS

MA participated in the conceptualization, data curation, formal analysis, software, investigation, methodology, project administration, resources, writing original draft, and writing review and editing of the draft manuscript. KEB, AM, RH, AA, MJN, JHV and BP involved in the conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, and writing review and editing of the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We acknowledge the study participants, the clinicians and laboratory technologists in the laboratory of the study site Hospitals for helping in the data collection. The study was funded by Armauer Hansen Research Institute, and Wellcome Trust Brighton and Sussex Centre for Global Health and Research.

CONSENT TO PUBLISH

Not applicable

References

1. Ethiopian Public Health Institute (EPHI). HIV Related Estimates and Projections for Ethiopia, March 2018, Addis Ababa, Ethiopia.
2. Abebe A, Kuiken CL, Goudsmit J, Valk M, Messele T, Sahlu T, et al. HIV type 1 subtype C in Addis Ababa, Ethiopia. AIDS Res Hum Retrovir. 1997; 13:1071–1075.
3. Hussein M, Abebe A, Pollakis G, Brouwer M, Petros B, Fontanet AL, et al. HIV–1 subtype C in commercial sex workers in Addis Ababa, Ethiopia. J Acquir Immune Defic Syndr. 2000; 23:120–127.
4. Abebe A, Pollakis G, Fontanet AL, Fisseha B, Tegbaru B, Kliphuis A, et al. Identification of a genetic subcluster of HIV type 1 subtype C (C') widespread in Ethiopia. AIDS Res Hum Retrovir. 2000; 16:1909–1914.
5. Dauwe K, Mortier V, Schauvliege M, Van Den Heuvel A, Fransen K, Servais JY, et al. Characteristics and spread to the native population of HIV–1 non-B subtypes in two European countries with high migration rate. BMC infectious diseases. 2015; 15:524.
6. Janssens W, Buve A, Nkengasong J. The Puzzle of HIV–1 Subtypes in Africa. AIDS. 1997; 11: 705–711.
7. Poss M, Gosink J, Thomas E, Kreiss J, Ndinya-Achola J, Mandaliya K, et al. AIDS Res Hum Retrovirol. 1997; 13: 493–499.
8. Hu DJ, Baggs J, Downing GR, Pieniazek D, Dorn J, Fridlund C, et al. Predominance of HIV–1 Subtype A and D Infections from Uganda. Emerging Infectious Diseases. 2000; 6: 609–615.

9. Louwagie J, Janssens W, Mascola J, Heyndrickx L, Hegerich P, van der Groen G, et al. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. J Virol. 1995; 69:263–271.
10. Salminen MO, Johansson B, Sonnerborg A, Ayehunie S, Gotte D, Lenikki P, et al. Full-Length Sequence of an Ethiopian Human Immunodeficiency Virus Type 1 (HIV–1) Isolates of the Genetic Subtype C. AIDS Res Hum Retroviruses. 1996; 12: 1329–1339.
11. Sherefa K, Johanssen B, Salminen M, Sonnerborrg A. Full Length Sequence of HIV–1 Subtype A Recombined with Subtype C in the Env Domain. AIDS Res Hum Retroviruses. 1998; 14:289–292.
12. Tscherning C, Alaeus A, Fredriksson R, Bjorndal A, Deng H, Littman DR, et al. Differences coreceptor Usage between Genetic Subtype of the HIV–1. Virology. 1998; 241: 181–188.
13. Abebe A, Demissie D, Goudsmit J, Brouwer M, Kuiken CL, Pollakis G, et al. HIV subtype C syncytium and non-syncytium Ethiopian patients with AIDS. AIDS. 1999; 13:1305–1311.
14. Moore JP, Kitchen SG, Pugach P, Zack JA. The CCR5 and CXCR4 co-receptors central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection. AIDS Res Hum Retroviruses. 2004; 20:111–126.
15. Nabatov AA, Pollakis G, Linnemann T, Kliphius A, Chalaby MI, Paxton WA. Intrapatient alterations in the human immunodeficiency virus type 1 gp120 V1V2 and V3 regions differentially modulate coreceptor usage, virus inhibition by CC/CXC chemokines, soluble CD4, and the b12 and 2G12 monoclonal antibodies. J Virol. 2004; 78:524–530.
16. Chowdhury IH, Potash MJ, Volsky DJ. Redefinition of Tropism of Common Macrophage-tropic Human Immunodeficiency Virus Type 1. AIDS Res Hum Retroviruses. 1995; 11: 1467–1471.
17. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R. Bioinformatics prediction of HIV coreceptor usage. Nat Biotechnol. 2007; 25:1407–1410.
18. Cashin K, Gray LR, Harvey KL, Perez-Bercoff D, Lee GQ, Sterjovski J, et al. Reliable genotypic tropism tests for the major HIV–1 subtypes. Scientific reports. 2015; 5:8543.
19. Adal M, Ayele W, Wolday D, Dagne K, Messele T, Tilahun T, et al. Evidence of genetic variability of human immunodeficiency virus type 1 in plasma and cervicovaginal lavage in ethiopian women seeking care for sexually transmitted infections. AIDS Res Hum Retroviruses. 2005; 21:649–653.
20. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, et al. Homozygous defect in HIV–1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV–1 infection. Cell. 1996; 86:367–377
21. Huang Y, Paxton WA, Wolinsky SM, Neumann AU, Zhang L, He T, et al. The role of a mutant CCR5 allele in HIV–1 transmission and disease progression. Nat Med. 1996; 2:1240–1243
22. de Roda Husman AM, Koot M, Cornelissen M, Keet IP, Brouwer M, Broersen SM, et al. Association between CCR5 genotype and the clinical course of HIV–1 infection. Ann Intern Med. 1997; 127:882–890.
23. Spudich SS, Ances BM. Neurologic Complications of HIV Infection: Highlights from the 2013 Conference on Retroviruses and Opportunistic Infections. Top Antivir Med. 2013; 21: 100–108.

24. Xu GG, Guo J, Wu Y. Chemokine Receptor CCR5 Antagonist Maraviroc: Medicinal Chemistry and Clinical Applications. Curr Top Med Chem. 2014; 14: 1504–1514.
25. Bonhoeffer S, Holmes EC, Nowak MA. Causes of HIV Diversity. Nature (Letter). 1995; 376: 125.
26. Mansky L. Retrovirus Mutation Rate and Their Role in Genetic Variatian. J Gen Virol. 1998; 79: 1337–1345.
27. Perelson AS, Neumann AU, Markowitz M. (1996). HIV–1 Dynamics In Vivo: Virion Clearance Rate, Infected Cell Life Span, and Viral Generation Time. Science. 271: 1582–1586.
28. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA. N-linked glycosylation of the HIV type–1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. J Biol Chem. 2001; 276:13433–13441.
29. Batra M, Tien PC, Shafer RW, Contag CH, Katzenstein DA. HIV type 1 Envelope Subtype C Sequences from Recent Seroconverters in Zimbabwe. AIDS Res Hum Retroviruses. 2000; 16: 973–979.
30. Lenz J, Su M, Mizrachi Y, Burke M, Rubinotein A. V3 Variation in HIV Seropositive Patients Receiving a V3 Targeted Vaccine. AIDS. 2001; 15: 577–581.
31. McMichael AJ. Triple bypass: complicated paths to HIV escape. J Exp Med. 2007; 204:2785–2787.
32. Adal M, Howe R, Kassa D, Aseffa A, Petros B. Associations of gender and serum total cholesterol with CD4+ T cell count and HIV RNA load in antiretroviral-naïve individuals in Addis Ababa. BMC Public Health. 2018a; 18:943.
33. Adal M, Howe R, Kassa D, Aseffa A, Petros B. Malnutrition and lipid abnormalities in antiretroviral naïve HIV-infected adults in Addis Ababa: a cross sectional study. PloS One. 2018b; 13:e0195942.
34. Gall A, Ferns B, Morris C, Watson S, Cotten M, Robinson M, et al. Universal Amplification, Next-Generation Sequencing, and Assembly of HIV–1 Genomes. J. Clin. Microbiol. 2012; 50:3838.
35. Hunt M, Gall A, Ong SH, Brener J, Ferns B, Goulder P, et al. IVA: accurate de novo assembly of RNA virus Genomes. Bioinformatics. 2015; 31: 2374–2376.
36. Salminen MO, Carr JK, Burke DS, Mccutchan FE. Identification of breakpoints in intergenotypic recombinants of HIV type–1 by bootscanning. AIDS Res Hum Retroviruses. 1995; 11:1423–1425.
37. Robertson DL, Sharp PM, McCutchan FE, Hahn BH. Recombination in HIV–1. Nature. 1995; 374:124–126.
38. Chesebro B, Wehrly K, Nishio J, Perryman S. Macrophage tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-trophic isolates: Definition of critical amino acids involved in cell tropism. J Virol. 1992; 66:6547–6554.
39. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis (MEGA) Version 6.0. Mol Biol Evol. 2013; 30: 2725–2729.
40. Gifford RJ, Liu TF, Rhee SY, Kiuchi M, Hue S, Pillay D, et al. The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV–1 drug resistance. Bioinformatics. 2009; 25: 1197–1198.

41. Kassu A, Fujino M, Matsuda M, Nishizawa M, Ota F, Sugiura W. Molecular epidemiology of HIV type 1 in treatment-naive patients in north Ethiopia. AIDS Res Hum Retrovir. 2007; 23: 564–568.
42. Mulu A, Maier M, Liebert UG. Deworming of intestinal helminths reduces HIV–1 subtype C viremia in chronically co-infected individuals. Int J Infect Dis. 2013; 17:e897–901.
43. Kalu AW, Telele NF, Gebreselasie S, Fekade D, Abdurahman S, Marrone G, et al. Prediction of coreceptor usage by five bioinformatics tools in a large Ethiopian HIV–1 subtype C cohort. PLoS ONE. 2017a. 12:e0182384.
44. Kalu AW, Telele NF, Gebreselasie S, Fekade D, Abdurahman S, Marrone G, et al. Monophylogenetic HIV–1C epidemic in Ethiopia is dominated by CCR5-tropic viruses-an analysis of a prospective country-wide cohort. BMC Infect Dis. 2017b; 17:37.
45. Bjorndal A, Sonnerborg A, Tschrning C, Albert J, Fenyo EM. Phenotypic characteristics of human immunodeficiency virus type 1 subtype C isolates of Ethiopian AIDS patients. AIDS Res Hum Retroviruses. 1999; 15:647–653.
46. Cilliers T, Nhlapo J, Coetzer M, Orlovic D, Ketas T, Olson WC, et al. The CCR5 and CXCR4 coreceptors are both used by human immunodeficiency virus type 1 primary isolates from subtype C. Journal of virology. 2003; 77:4449–4456.
47. Neogi U, Prarthana SB, D'Souza G, Decosta A, Kuttiatt VS, Ranga U, et al. Co-receptor tropism prediction among 1045 Indian HIV–1 subtype C sequences: Therapeutic implications for India. AIDS research and therapy. 2010; 7:24.
48. Ferbas J, Giorgi JV, Amini S, Grovit-Febas K, Wiley DJ, Detus R, et al. Antigen Specific Production of RANTES, Macrophage Inflammatory Protein (MIP)–1 and MIP–1 In Vitro is a Correlate of Reduced Human Immunodeficiency Virus Burden In Vivo. J Infect Dis. 2000; 182:1247–1250.
49. Ceresola ER, Nozza S, Sampaolo M, Pignataro AR, Saita D, Ferrarese R, et al. Performance of commonly used genotypic assays and comparison with phenotypic assays of HIV–1 coreceptor tropism in acutely HIV–1-infected patients. The Journal of antimicrobial chemotherapy. 2015; 70:1391–1395.
50. Overbaugh J, Kreiss J, Poss M, Lewis P, Mostad S, John G, et al. Studies of human immunodeficiency virus type 1 mucosal shedding and transmission in Kenya. J Infect Dis. 1999; 179: S401–404.

Figures

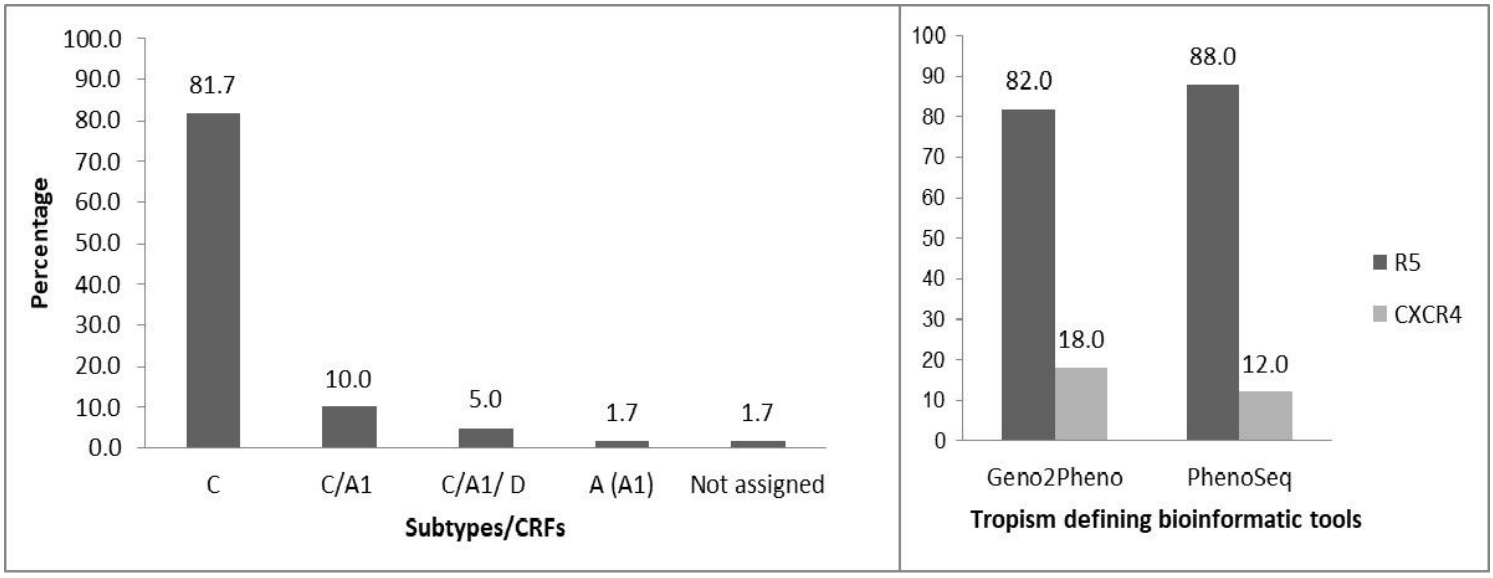

Figure 1

Percentage of subtypes, circulating recombinant forms, and R5 and CXCR4 coreceptor uses.

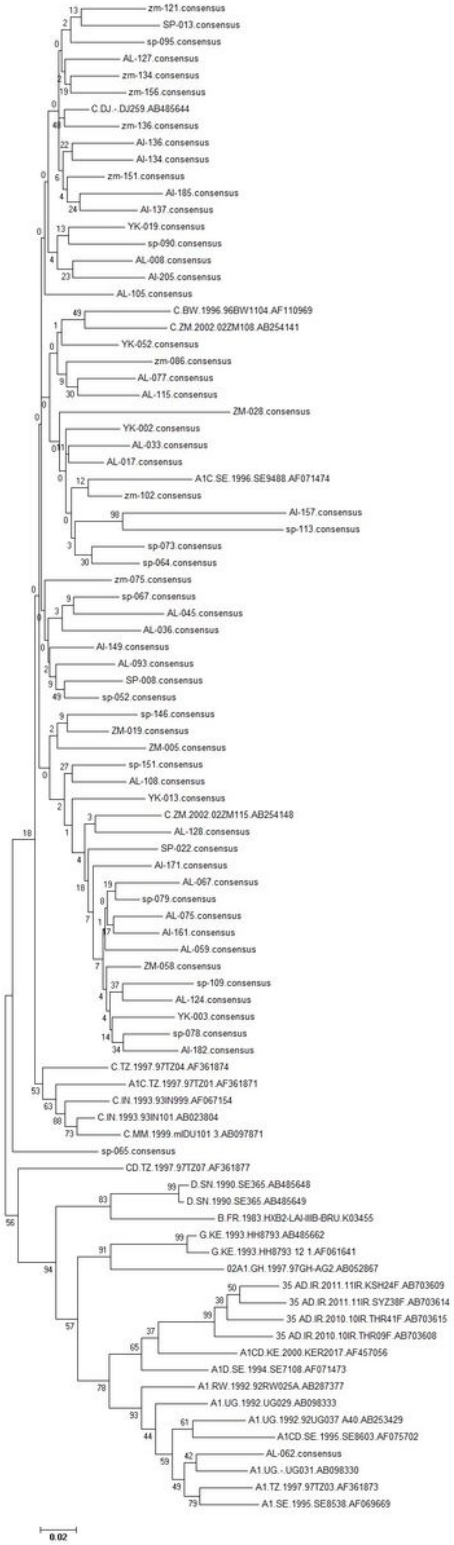

Figure 2

Neighbor-joining tree demonstrating the evolutionary relationship and the distance of the HIV-1 genome consensus sequences. Sixty sequences from plasma samples, subtype C, Subtupe A1, Subtype D and AC, AD, CD, ACD circulating recombinant forms as reference sequences from the Los Alamos database were used. The scale bar represents a genetic distance of 2%.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- S1Fig.pdf