AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 16, Number 17, 2000, pp. 1909–1914
Mary Ann Liebert, Inc.

Sequence Note

Identification of a Genetic Subcluster of HIV Type 1 Subtype C (C′) Widespread in Ethiopia

ALMAZ ABEBE,[1] GEORGIOS POLLAKIS,[2] ARNAUD L. FONTANET,[3] BITEW FISSEHA,[4] BELETE TEGBARU,[4] ALETTA KLIPHUIS,[2] GIRMA TESFAYE,[5] HAILU NEGASSA,[6] MARION CORNELISSEN,[2] JAAP GOUDSMIT,[2] and TOBIAS F. RINKE DE WIT[1]

ABSTRACT

Others and we have previously shown that subtype C is the predominant HIV-1 subtype and the major cause of AIDS in Ethiopia. The present study shows that subtype C in Ethiopia has a genetic subcluster, designated C′, has not increased in frequency, or spread geographically, over the period 1988 (%C′ = 23/53) to 1996–1997 (%C′ = 26/50). There is no association of the HIV-1 subtype C or subcluster C′ with geographic location, time of sample collection, or risk group in Ethiopia. Of 105 randomly collected samples representing 7 different towns in Ethiopia, all but 2 (1 subtype A from Addis Ababa, 1997 and 1 subtype D from Dessie, 1996) belong to subtype C.

HUMAN IMMUNODEFICIENCY VIRUS type 1 (HIV-1) subtypes are distributed unevenly across Africa nations.[1–3] In East and Central African countries such as Uganda, Kenya, and Tanzania, the HIV-1 epidemic involves mainly two HIV-1 subtypes, A and D.[4,5] In contrast, subtype C has dominated the rapidly expanding epidemic in Botswana, South Africa, and Ethiopia.[6–8] HIV-1 subtype C is on the rise, possibly gradually replacing subtype D viruses in East African countries.[9]

The relative roles played by virological, behavioral, and host determinants in the epidemic expansion of any particular HIV-1 subtype are unclear. Careful surveillance of genetic subtypes in a given population is presently a particularly important approach for better understanding the biological properties of different HIV-1 subtypes.

The first Ethiopian HIV-1-positive sera were detected in 1984[10] and the first Ethiopian AIDS cases were reported in 1986 in Addis Ababa, the capital city, with presently more than 3 million inhabitants.[11] A national surveillance performed in 1988 among commercial sex workers in 23 towns and cities in Ethiopia revealed an HIV-1 prevalence of 1 to 38%.[12] In 1994 the seroprevalence was 7% among blood donors.[13] Sentinel surveillances performed in 1995, 1996, and 1997 in Addis Ababa report HIV-1 prevalence ranging from 14 to 20% in pregnant women.[13] Studies performed in 1997 and 1998 reveal an HIV-1 prevalence of 45–74% in commercial sex workers in Addis Ababa (M. Aklilu, personal communication; ref. 14). Finally, sequence data on sera and plasma samples collected in Addis Ababa demonstrated that the majority of the circulating viruses belong to subtype C.[14–16] These Ethiopian subtype C sequences differ slightly from the consensus C sequence and there was some hint of the presence of a separate subcluster within the main C group. Nevertheless, the presence of such a subcluster was not supported by a significant bootstrap value in phylogenetic tree analysis.

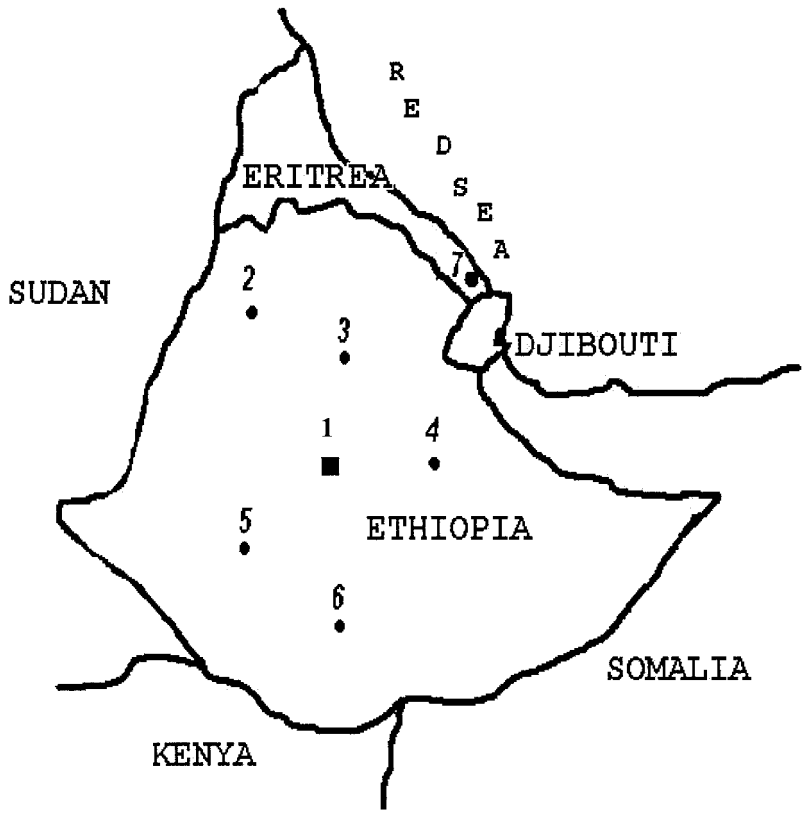
To assess the geographical distribution in more detail and the possible influx of HIV-1 subtype(s) in the whole of Ethiopia,

[1]Ethiopian–Netherlands AIDS Research Project (ENARP) at the Ethiopian Health and Nutrition Research Institute (EHNRI), Addis Ababa, Ethiopia.
[2]Department of Human Retrovirology, Academic Medical Center, University of Amsterdam, 1105 AZ, Amsterdam, The Netherlands.
[3]Division of Public Health and Environment, Municipal Health Service, 1000 CE, Amsterdam, The Netherlands.
[4]National Referral Laboratory for AIDS, EHNRI, Addis Ababa, Ethiopia.
[5]Ethiopian Red Cross Society-National Blood Transfusion Service, Addis Ababa, Ethiopia.
[6]AIDS/STD Control Team, Ministry of Health, Addis Ababa, Ethiopia.

105 serum samples that have been collected from seven different towns and two risk groups were analyzed. Figure 1 shows the map of Ethiopia, the number of serum samples sequenced, the towns, and the year of sample collection.

The procedures of HIV-1 RNA isolation, reverse transcription, and direct sequencing were described earlier.[15] The C2–V3 region of the envelope gp120 glycoprotein gene was sequenced and sequence alignment was performed manually according to the Los Alamos reference sequences used for subtyping. The nucleotide alignments were subjected to phylogenetic tree analysis, using neighbor-joining and maximum likelihood methods implemented in the PHYLIP package programs[17] and using Kimura two-parameter distances. The bootstrap option in the MEGA program was used to determine the reliability of the clusters in the phylogenetic trees. A bootstrap value equal to or greater than 75% was considered significant.[18] The synonymous and nonsynonymous nucleotide substitution distance matrix was generated by the MEGA program, according to the Jukes and Cantor method.[19] Multivariate principal coordinate analysis was done with PCOORD software[20] and amino acid sequence comparisons were done with the VESPA program.[21]

The phylogenetic analysis depicted in Fig. 2 clearly shows that subtype C is widely distributed and dominates the HIV-1 epidemic in Ethiopia. Among the 105 samples, 2 sequences clustered



Town	Number of samples collected 1988(C')	1996/7(C')
1. Addis Ababa	7(4)	11(6)
2. Gondar	7(5)	9(3)
3. Dessie	7(1)	5(2)
4. Dire Dawa	9(2)	8(3)
5. Jimma	7(2)	10(7)
6. Arba Minch	9(7)	7(5)
7. Assab	7(2)	0
Total	53(23)	50(26)

FIG. 1. Map of Ethiopia and Eritrea. The numbers in the map indicate the different towns included in the study. The samples shown in the table were collected from commercial sex workers in 1988 and blood donors in 1996–1997. The numbers in parentheses indicate the number of isolates that belong to the subcluster C′.

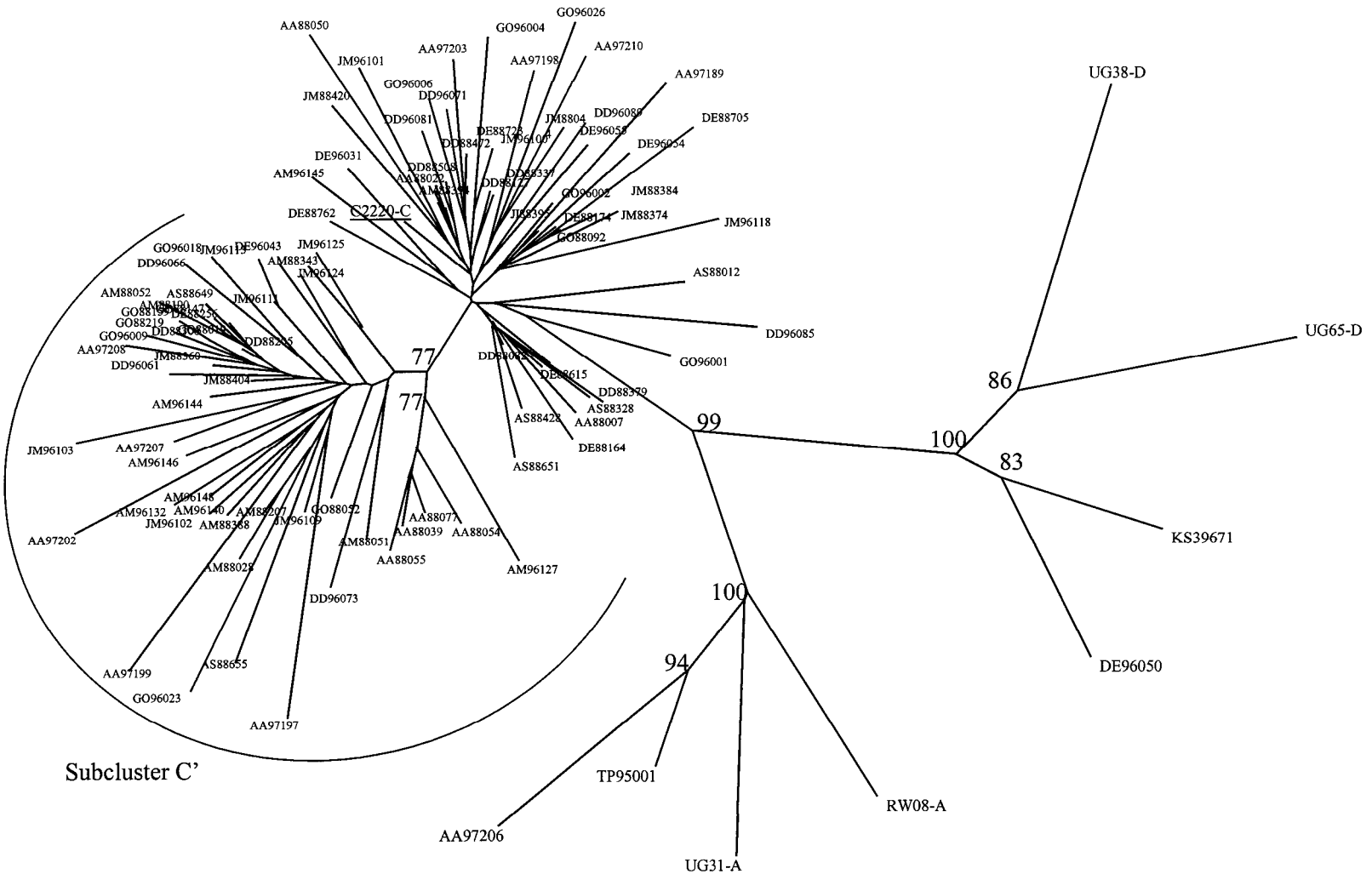


FIG. 2. Phylogenetic tree analysis of Ethiopian HIV-1 sequences by neighbor-joining algorithm.[17] The subcluster is indicated as subcluster C′ in the tree. The HIV-1 subtype A (AA97206) and subtype D (DE96050) isolates found in this study are shown, clustering with the subtype A (TP95001) and the subtype D (KS39671) from our previous studies.[14,16] Sequences have been indicated by codes: AA, Addis Ababa; AM, Arba Minch; AS, Assab; DD, Dire Dawa; DE, Dessie; JM, Jimma; GO, Gondar. The first two-digit number following these codes indicates the year of sample collection, and the next three digits indicate sample number. The GenBank accession numbers are AF245518 to AF2456613. Numbers by the branches represent bootstrap values out of 100 replications.

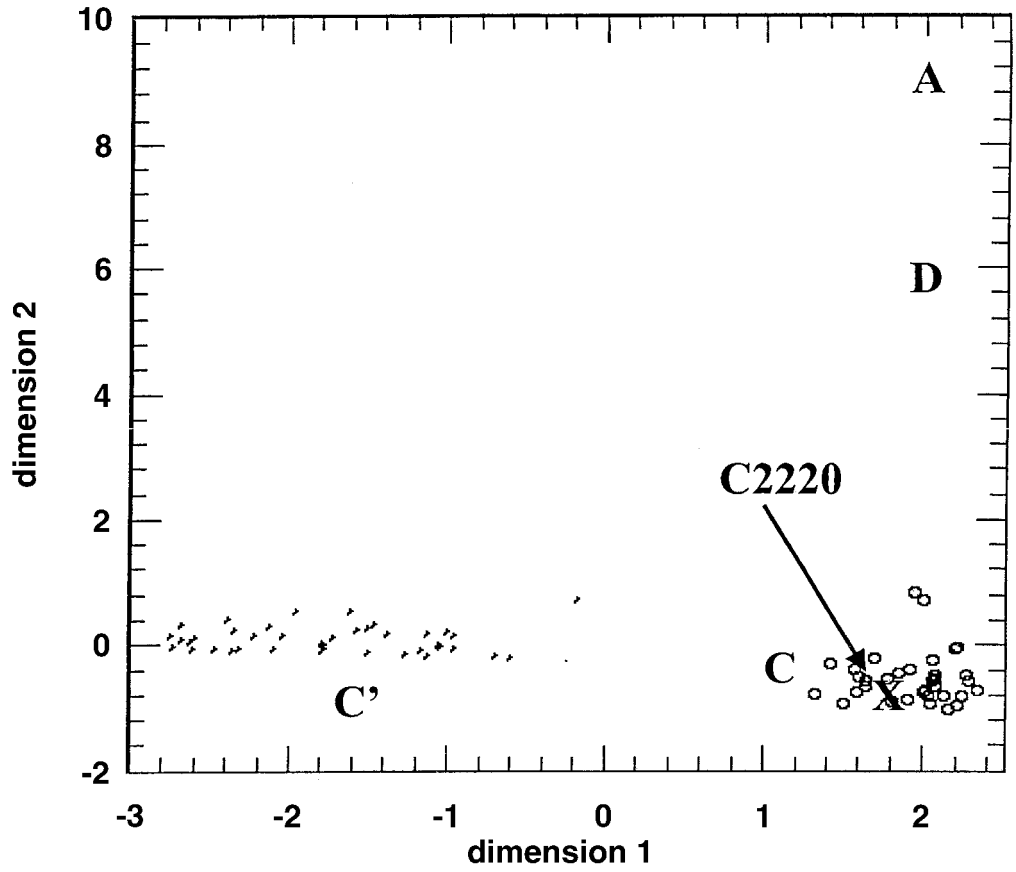


FIG. 3. Result of the multivariate principal coordinate analysis (PCOORD) for the main group C and the subcluster C′. Axes are the two dimensions that were first extracted; together they cover 24% of the total difference between the two groups. The first 10 axes cover 53% of variation. The arrow points the subtype C reference C2220. Two reference sequences of subtypes A and D were included, noted as such.

with previously published subtype A and subtype D sequences from Addis Ababa.[14,15] Sample AA97026 of Addis Ababa showed homology with TP95001 (sutype A, pregnant women, Addis Ababa, 1995),while DE96050 from Dessie clustered with KS39671 (subtype D, commercial sex worker, Addis Ababa, 1997). All serum samples collected in 1988 contained subtype C virus. Although both HIV-1 subtype A and D isolates were collected in 1996–1997, there is evidence of an earlier introduction of HIV-1 subtype A in Ethiopia, in or even before 1991.[22]

Within the Ethiopian sequences a subcluster (named C′) could be identified with a significant bootstrap value (bootstrap 77%). The Ethiopian isolate C2220 from 1986, used as a C reference, does not belong to the C′ subcluster. The prevalence of both the subcluster C′ and the main group C viruses appears equally distributed, as this study indicates that 48% of the analyzed serum samples contain a C′ virus and 52% contain a C virus. The subcluster was confirmed by maximum likelihood phylogenetic analysis (data not shown) and by multivariate principal coordinate analysis (Fig. 3).

The Ethiopian subcluster C′ viruses are grouped together, independent of any variable considered in this study, which includes geography, risk group, and time of sample collection. The two groups cocirculate with similar prevalence, although the rate of synonymous and nonsynonymous nucleotide variation among the sequences of the C group is higher than in the C′ group. These differences are statistically significant ($p < 0.0001$) for the two time points of sample collection.

VESPA software was used for the amino acid sequence comparison of the Los Alamos database[3] subtype C sequences (LsA C) and the two subgroups (Eth C) and C′ (Eth C′) circulating in Ethiopia. Both groups C and C′ show significant amino acid differences when compared with the database or with each other, but the main group C is genetically closer to the database sequences than to the subcluster C′ sequences (Fig. 4). The presence in the C′ group of a lysine (K) at position 304 instead of a glutamic acid (E) affects the number of positive charges in the V3 loop and a valine (V) instead of an asparagine (N) at position 294 leads to the loss of a potential N-glycosylation site. We do not have evidence that there are significant biological differences between the groups C and C′, although it is known that the number of charges and glycosylation in the V3 loop can affect cellular tropism and neutralization ability of antibodies.[23,24] Further studies and follow-up of the epidemic are needed to answer these questions.

We conclude that viruses circulating in Ethiopia since 1988 cluster with the main subtype C, but a significant subcluster C′

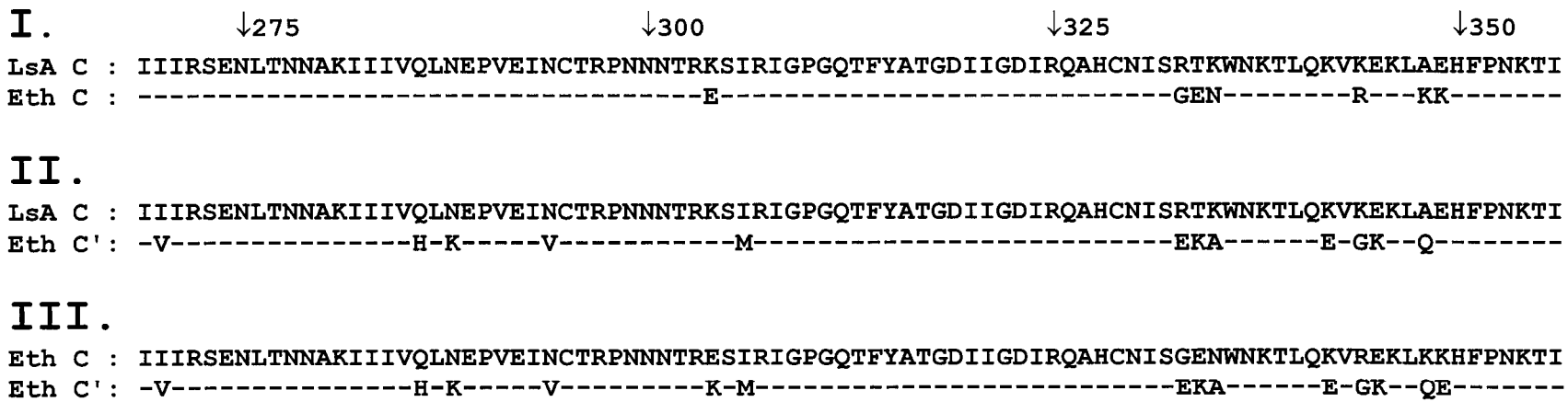


FIG. 4. VESPA-supported amino acid sequence comparison between the Los Alamos database subtype C (LsA C) sequences and the C (I) and C′ (II) groups in Ethiopia, indicated as Eth C and Eth C′, respectively.

was noted in multiple analyses. This subcluster of subtype C (C′) has been in a fifty-fifty equilibrium with the main subtype C in Ethiopia since 1988.

ACKNOWLEDGMENTS

This study is part of the Ethiopian–Netherlands AIDS Research Project (ENARP), a collaborative effort of the Ethiopian Health and Nutrition Research Institute (EHNRI), the Amsterdam Municipal Health Service (GG/GD), the Central Laboratory of the Netherlands Red Cross Blood Transfusion Service (CLB), and the Academic Medical Center of the University of Amsterdam (AMC). ENARP is financially supported by the Netherlands Ministry of Foreign Affairs and the Ethiopian Ministry of Health (MOH) as a bilateral project. The authors thank Mr. Tesfaye Mebratu for voluntary technical assistance and blood bank technicians of the various towns for collecting and sending the samples included in the study.

REFERENCES

1. Janssens W, Buve A, and Nkengason JN: The puzzle of HIV-1 subtypes in Africa. AIDS 1997;11:705–712.
2. Lasky M, Perret J-L, Peeters M, Binbollet-Ruche F, Liegeois F, Patrel D, Molinier S, Gras C, and Delaporte E: Presence of multiple non-B subtypes and divergent subtype B strains of HIV-1 in individuals infected after overseas deployment. AIDS 1997;11: 43–51.
3. Korber B, Kuiken C, Foley BT, Hahn BH, Mccutchan FE, Mellors J, and Sodroski J: *Human Retroviruses and AIDS 1998: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences.* Los Alamos National Laboratory, Los Alamos, New Mexico, 1999.
4. Poss M, Gosink J, Thomas E, Kreiss JK, Ndinya-Achola J, Mandaliya K, Bwayo J, and Overbaugh J: Phylogenetic evaluation of Kenyan HIV type 1 isolates. AIDS Res Hum Retroviruses 1997; 13:493–499.
5. Robbins KE, Bandea CI, Levin A, Goedert JJ, Blattner WA, Brubaker G, Brown TM, Schochetman G, Kalish ML, Shao J, and O'Brien TR: Genetic variability of human immunodeficiency virus type 1 in rural northwest Tanzania. AIDS Res Hum Retroviruses 1996;12:1389–1391.
6. Novitsky VA, Montano MA, McLane MF, Renjifo B, vannberg F, Foley BT, Ndun'u TP, Rahman M, Makhema MJ, Marlink R, and Essex M: Molecular cloning and phylognetic analysis of human immunodeficiency virus type 1 subtype C: A set of 23 full length clones from Botswana. J Virol 1999;73:4427-4432.
7. van Harmelen JH, van der Ryst E, Loubser AS, York D, Madurai S, Lyons S, Wood R, and Williamson C: A predominantly HIV type 1 subtype C-restricted epidemic in South African urban populations. AIDS Res Hum Retroviruses 1999;15:395–398.
8. Sherefa K, Sallberg M, and Sonnerborg A: Evidence of no change in V3 loop antibody recognition pattern in HIV type 1-infected Ethiopians between 1988 and 1993. AIDS Res Hum Retroviruses 1994;11:1551–1555.
9. Renjifo B, Chaplin B, Mwakagile D, Shah P, vannberg F, Msamanga G, Hunter D, Fawzi W, and Essex M: Epidemic expansion of HIV type 1 subtype C and recombinant genotypes in Tanzania. AIDS Res Hum Retroviruses 1998;14:635-638.
10. Tsega E, Mengesha B, Nordenfelt E, Hansson B-G, and Lindberg J: Serological survey of human immunodeficiency virus infection in Ethiopia. Ethiop Med J 1988;26:179–184.
11. Lester FT, Ayehunie S, and Zewdie D: Acquired immunodeficiency syndrome: Seven cases in Addis Ababa hospital. Ethiop Med J 1988;26:139–145.
12. Mehret M, Khodakevich L, Zewdie D, Ayehunie S, Gizaw G, Shanko B, Manore H, Gemeda A, Ketema F, Yadeta M, Bekele D, Kebede T, Fessehaye T, Lakew S, and Bekele R: HIV-1 infection and related risk factors among female sex workers in urban areas of Ethiopia. Ethiop J Health Dev 1990;4:163–170.
13. Fontanet AL, Messele T, Dejene A, Enquselassie F, Abebe A, Cutts TF, Rinke de Wit TF, Sahlu T, Bindels P, Yeneneh H, Coutinho AR, and Nokes JD: Age- and sex-specific HIV-1 prevalence in the urban community setting of Addis Ababa, Ethiopia. AIDS 1998;12:315–322.
14. Hussein M, Abebe A, Pollakis G, Brouwer M, Petros B, Fontanet AL, and de Wit TF: HIV-1 subtype C in commercial sex workers in Addis Ababa, Ethiopia. J Acquir Immune Defic Syndr 2000;23:120–127.
15. Abebe A, Kuiken CL, Goudsmit J, Valk M, Messele T, Sahlu T, Yeneneh H, Fontanet AL, de Wolf F, and Rinke de Wit TF: HIV type 1 subtype C in Addis Ababa, Ethiopia. AIDS Res Hum Retroviruses 1997;13:1071–1075.
16. Abebe A, Demissie D, Goudsmit J, Brouwer M, Kuiken CL, Pollakis G, Schuitemaker H, Fontanet AL, and Rinke de Wit TF: HIV-1 subtype C syncytium and non-syncytium inducing phenotypes and coreceptor usage among Ethiopian patients with AIDS. AIDS 1999;13:1305–1311.
17. Feelsenstein J: http://evolution.genetics.washington.edu/phylip.html. *PHYLIP: Phylogeny Inference Package,* version 3.52. University of Washington, Seattle, Washington, 1996.
18. Hillis DM and Bull JJ: An empirical test for bootstrapping as a method for assessing confidence in phylogenetic analysis. System Biol 1993;42:182–192.
19. Kumar S, Tamura K, and Nei M: *Molecular Evolutionary Genetics Analysis (MEGA),* version 1.01. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, Pennsylvania, 1993.
20. Higgins DG: Sequence ordinations: A multivariate analysis approach to analysing large sequence data sets. Comput Appl Biosci 1992;8;15–22.
21. Korber B and Myers G: Signature pattern analysis: A method for assessing viral sequence relatedness. AIDS Res Hum Retroviruses 1992;9:1549–1560.
22. Sherefa K, Johansson B, Salminen M, and Sonnerborg A: Full-length sequence of human immunodeficiency virus type 1 subtype A, recombined with subtype C in the env.V3 domain. AIDS Res Hum Retroviruses 1998;14:289–292.
23. Back NK, Smit L, de Jong JJ, Keulen W, Schutten M, Goudsmit J, and Tersmette M: An N-glycan within the human immunodeficiency virus type 1 gp120 V3 loop affects virus neutralization. Virology 1994;199:431-438.
24. de Jong JJ, de Ronde A, Keulen W, Tersmette M, and Goudsmit J: Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution. J Virol 1992;66: 6777–6780.

Address reprint requests to:
Almaz Abebe
EHNRI
Ras Desta Dantwe Street
P.O. Box 1242
Addis Ababa, Ethiopia

E-mail: enarp@telecom.net.et