AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 13, Number 12, 1997
Mary Ann Liebert, Inc.

Sequence Note

HIV Type 1 Subtype C in Addis Ababa, Ethiopia

ALMAZ ABEBE,[1] CARLA L. KUIKEN,[2] JAAP GOUDSMIT,[2] MARGREETH VALK,[2] TSEHAYNESH MESSELE,[1] TEFERA SAHLU,[1] HAILU YENENEH,[1] ARNAUD FONTANET,[1] FRANK DE WOLF,[2] and TOBIAS F. RINKE DE WIT[1]

HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 variants in different geographic regions have been phylogenetically classified into several distinct genetic subtypes (A–I and O) on the basis of sequence differences in the V3 regions of their gp120 envelope genes.[1,2] In Africa the existence of all genetic subtypes except subtype I has been confirmed. The presence of this multitude of subtypes indicates the extensive divergence HIV-1 has accumulated in the African continent.[3–5] This study describes the distribution of HIV-1 subtypes in Ethiopia, East Africa. The first Ethiopian AIDS case was reported in 1986 and the AIDS epidemic is a rapidly growing problem in Addis Ababa, the capital city with more than 2 million inhabitants.[6] In Addis Ababa the HIV-1 seroprevalence is estimated to be 10–27% in pregnant women (PW), 47–59% in commercial sex workers (CSWs) (ENARP sentinel surveys 1995 and 1996) and 7% among blood donors (BDs) (Ethiopian Red Cross Society, National Blood Transfusion Service [ERCS-NBTS], unpublished data, 1994). In the countries neighboring Ethiopia subtypes A (Djibouti, Kenya), C (Kenya, Djibouti, Somalia), and D (Kenya) are present.[1,7] Preliminary sequence data on a limited number of samples have indicated the presence of subtype C in Addis Ababa in 1988.[8] To assess the distribution of HIV-1 subtypes in Addis Ababa in more detail, 94 sera were analyzed, collected from 3 different risk groups over 1989–1995 (Table 1).

HIV-1 RNA was isolated from the collected sera according to a standard procedure.[9] Viral RNA was converted to cDNA and then subjected to a nested polymerase chain reaction (PCR) amplifying a 284-bp fragment covering the V3 region of the gp120 gene (positions 785–1069, HIV-1 subtype B consensus database, Los Alamos, 1993). First and second PCR conditions were as follows: 35 and 25 cycles of, respectively, 1 min at 95°C, 1 min at 55°C, and 2 min at 72°C, followed by a final incubation for 10 min at 72°C. The amplified products were sequenced directly by automated cycle sequencing with dye terminators, using an ABI (Foster City, CA) 370 A system. Sequence alignments and comparisons were performed by using CLUSTAL and the neighbor-joining algorithm.[10,11] Figure 1 demonstrates the results of neighbor-joining sequence comparisons, clearly indicating that the vast majority of sequenced samples (93 of 94) are of the C subtype. The prevalences of C and A subtypes among the 57 samples analyzed from 1995 were, respectively, 56 of 57 (98.2%) and 1 of 57 (1.8%). In addition, our data show that the C subtype is highly abundant in samples collected before 1995.

The Ethiopian subtype C sequences tend to differ slightly from the consensus C sequence (HIV-1 subtype C consensus database, Los Alamos, 1995). Within the Ethiopian sequences, a subcluster could be identified that is fairly homogeneous. Illustrating this homogeneity, pairwise nucleotide comparisons revealed a 7.4% difference within this subcluster, as compared to an 11.5% difference within the remaining sequences. The latter are more related to HIV-1 subtype C sequences from different countries (Fig. 1). Statistical comparisons, using a previously published method,[12] revealed significant sequence differences between the main group and the subcluster at positions 270, 272, 278, 286, and 294 (Fig. 2). The presence of a subcluster of the C subtype may suggest an independent introduction of the HIV subtype C virus in Addis Ababa. However, the subcluster could not be detected with significant preference in

TABLE 1. STUDY SUBJECTS, NUMBER OF SERA SEQUENCED, AND YEAR OF SAMPLE COLLECTION

Year	*Subjects*	*Samples tested*
1989	Commercial sex workers	7
1990	Commercial sex workers	6
1992	Commercial sex workers	24
1995	Commercial sex workers	19
1995	Pregnant women	9
1995	Blood donors	29
	Total:	94

[1]Ethiopian–Netherlands AIDS Research Project (ENARP), Ethiopian Health and Nutrition Research Institute (EHNRI), Addis Ababa, Ethiopia.
[2]Human Retrovirus Laboratory, Department of Human Retrovirology, University of Amsterdam, 1105 AZ Amsterdam, The Netherlands.

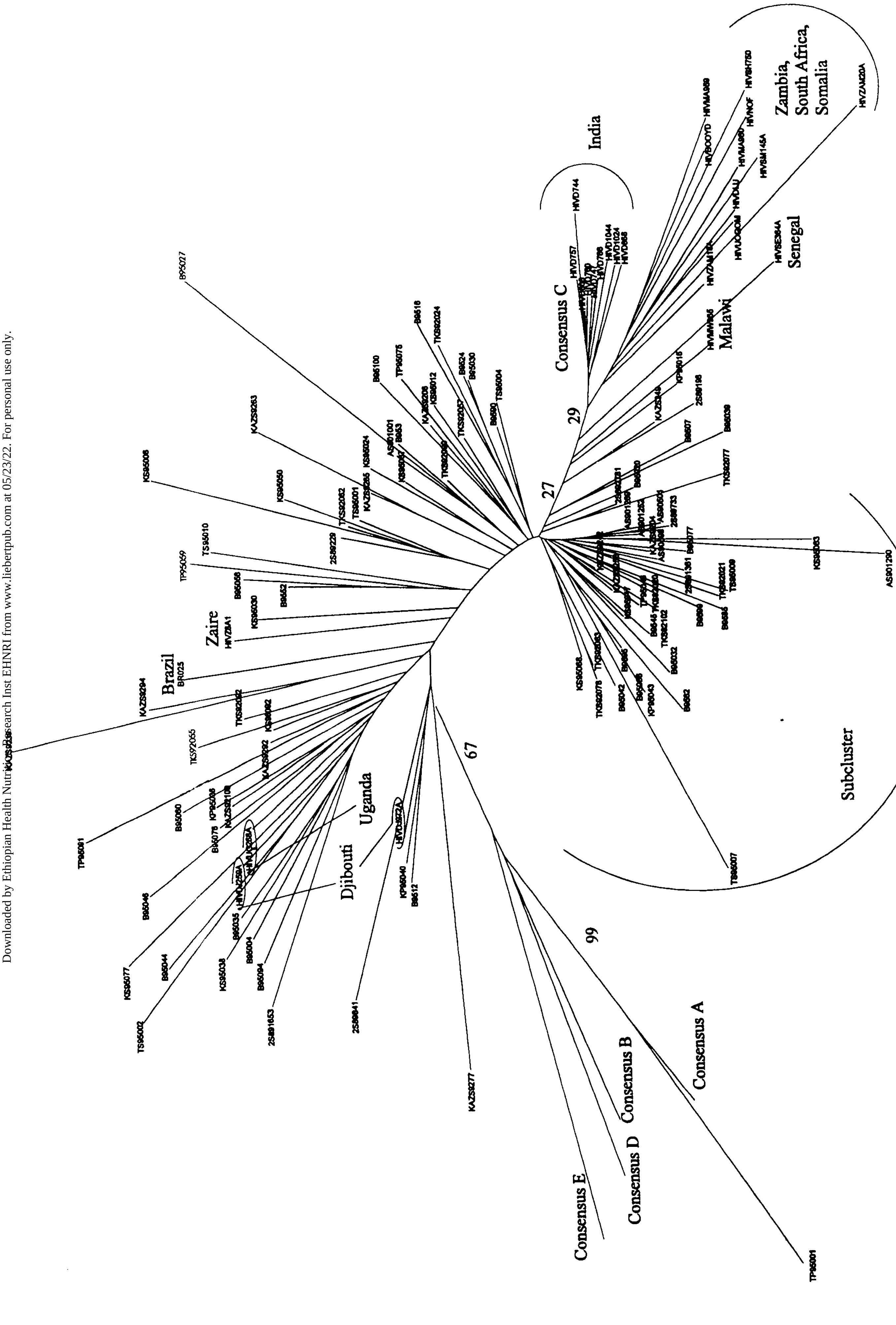

FIG. 1. Phylogenetic tree analysis of 94 Ethiopian HIV-1 sequences by neighbor-joining algorithm.[17] The main group and a subcluster are indicated; in addition, consensus sequences for HIV-1 A–E subtypes and different C-subtype sequences from various geographical areas are shown. Sequence names have been indicated by codes: 2S89 (. . .), CSWs 1989; AS90 (. . .), CSWs 1990; KAZS92 (. . .), CSWs 1992; TKS92 (. . .), CSWs 1992; TP95 (. . .), PW 1995; KP95 (. . .), PW 1995; KS95 (. . .), CSWs 1995; TS95 (. . .), CSWs 1995; B95 (. . .), BD 1995. Numbers by the branches represent bootstrap values out of a 100 replications.

```
Main Group
               270                 290                 310                 330                 350
                |                   |                   |                   |                   |
                * *     *       *       *
Consensus:     IVIRSENLTNNAKIIIVQLKEPVEIVCTRPNNNTRKSMRIGPGQTFYATGDIIGDIRQAHCNISEETWNKTLQEVGKKLQEHFPNKTI
A_TP95001      -R-----I-----N----FSD--N-T-----------V------------------------V-RSA------Q-VIQ-KT-VG----
C_2S891653     TI------------------D--Q-N---------E-I--------F---G-------------RAD-----E--KG--KGL-N--N-
C_2S89229      -I-----------------------N--------------------------T-----------KA----------E--RK-----N-
C_2S892781     -----------V-L---H--N-----------------------------A--------------K---D-------E----------
C_AS901001     -I------------------N-----------------------------G--------Y-S---K---D-----SEE-R--------
C_B9507        -------------T---H--DH---E--------------------------------------KDK----------Q--K-------
C_B9516        -------------------NQS----------------------------E---N----Y---N-KA--E---N-S-------T----
C_B9512        TI-----------------NKS---N-----------I------------E-------------RQK--E--ER-KE--KGY------
C_B953         -I-----------------NN-I---------------------A-----A--------Y----------------EE-X-----XS-
C_B9524        -------------T------K--------------R-T-------------------K-Q------------HD-S---L--------
C_B9552        -T---------V--------D----N--------------------------------------G----------KR---K-------
C_B9590        ---K---R---X-X---X---H---T----X-----X---------------------------GK----------R-----------
C_KAZS9249     -I---------V-----X-NNSX-----------------------------------------R-R--------R---X--------
C_KP95016      -----------V-T---H-N-S-------------R--------------A--------------GK--A---K-EN--------R--
C_KS95024      -I-------------------S-X------------------------------N----Y-----KN--D---K-----L-Y----S-
C_KP95040      -I-----------------NQS---N-----------I------------A-------------RQQ-----DK-RT--K---.-T--
C_KS95006      -I-----------------------N---------R-I------ALFHAT-----------T-N-T--HN---K---------.----
C_KS95029      TI-----------T-----N-T---N----G---------------------------------G-N------K-KE--G-L------
C_KS95050      -I--------------I---K--N-T-E---------V---------------------------RE--N---K-----D--------
C_TP95091      TI----------------FN-T---N-K---------F----------DI-GE-----------G-K--A---R-SR--G-----R--
C_TS95002      -I-----I-----------N-S-A-N-----------IG------L----A-----------V-G-K--V--ER-KN--T-Y-N-SIK
C_TS95010      -I------A----T---------K-------------V---------------------Y-H-NGT------E--KG--RK--L----
C_TS95004      -----------V-T---H-TK--------------------------------------------K---------RG----R------
C_TS95001      T--K-K--------------K--Q-N----------------------------------------A---------------------
C_B95094       VI-----------------N-T-Q-N------T--R-IG-------F------------X----GDQ-XT--E--XE--A---.-R--
C_TP95059      V------------T--I------Q-N-----------V------A-F---G-------------GQK-----ED-K---K---.----
C_TKS92062     -I-----------------------N---------------------------------------KA--D---K--E--K--------
C_KS95077      DI------A-----------N--Q-N----G----E-IG-----A----------T---Y----G-E--E--HR--E--IRL-SKNIT
C_KS95030      -I-----------T-----------N----G----R-G-----R-------V-------Y----RTK--------------Q------
C_B95100       -------------V---H------------S------V------------G-T------------K---S---V-----L------NT
C_B95056       -T---------V--------D----N---------E----------------------------GGQ------G-KE--K--------
C_B95046       -I-----------------N-S-P-T---------E-I------------A--------Y----GWDAT..--K-KGE-AK----R--
C_B95020       ---------D-V-----H-HD--------------------------------------------KE-------------K-------
C_B95027       -I---------D--------TT-------R---------------L-E--E--------Y-----K---D--RD--E--G-T-Y-IXX
C_2S892841     -----------------H-N-S---N-----------I-------L----G--------Y----G-K-----D---PE-RK-SIKHKX
C_2S891896     ---------D-V-----H-N-S------------------------------------------G-R--T---R-RR--E---H-R--
C_B95004       -I-----------T-----NHT---N-----------I--------F--DA-------------RAQ--T--ER-RE--RK--N-T-T
C_KAZS9239     -I---------V-T-----RNH---N-----------IG----------DE--------------GK-STV-RK-ER------LMHSN
C_KAZS9265     -----------V-T---H-------------------V----------------N----Y-----KA--D------G----S---R--
C_KAZS92109    TI--A--------------N-----N---------E-I------------S-------------R-K-KN---K-EER-KK-------
C_KP95036      -I---------V-T-----N-T---N---------E-I--------------------------R-K--N--ER-KER-RK---SR--
C_KS95012      ------S----------H-TK---------S---------------------T------------K---D----------E-----S-
C_KS95038      -I-----------T-----N-T-Q-N----S------I-L----A-----G---E---------RDK-----EK-KE--KR--NGTIV
C_TKS92055     -MV--------V-------N-T---N---------R-I----------------N----Y----G-K--S---K-K---E-L-H----
C_TKS92077     -----------V-----H-Q-----E-------------V---------------------I---KK--D---K--SE----------
C_TKS92092     -I------I----T---H-N-S---N-----------L-------------------K-Y------K-KQ---R-SA--E--------
C_TP95075      -----K---D-V------------------G------V---------------------Y-----K--SD--RD-S--------K---
C_TKS92090     TI-------D-V--------N-----------------------------A-----------------------------------S-
C_TKS92057     -----------V------------------S--------------------------------------N------ER--K-------
C_TKS92024     -----------------H-T-A--------G--------------------------K-----NRQ---E---N-SR-----------
C_KAZS9277     VM-----------------------N-----------I--------F---EV--N----Y--V-RQA--Q-----A-Q-RK--N----
C_KAZS9263     -I---------V-T-----MK----N-----------IG---------------N----Y-----KA--D---K--R----S---R--
C_KAZS9206     ---------D-V-V-----NK----------------VG--------------------Y-----Q---E---G------------N-
C_KAZS9292     TI-----------------N-S---T-----------I-------------------------NG-K--A---K-RE--EK--L----
C_KAZS9294     -I-----------------N-T---N----G------I------A-----E------K-------GK-RTV--K-E---------G--
C_KS95087      -I--F--------------------------------V--------------V-N-----------K------------X------SK
C_B95060       TI----D-A----T-----DK----N---------E-I------------G-------------A-G--Q---K-TV--A--------
C_B95039       ---------D-V-T---H--D----R------------------------G--------Y---N-SK--X-------E-GK-------
C_B95030       -------------T-----------K-----------T-------------------K--------------HD-S---R-----R--
C_B95035       -I-----------------N-T---N-----------I------------A-------------R-K-TR---R-KE--R---NKTME
C_B95076       -I-----------------N-T-Q-N---------E-I--------------------------GR--KA--EK-KD--KK----GE-
C_B95044       -I-----------------NKS---N----ST---E-I------------------------V-G-K--A--KK-VN--KK--D--RP
```

FIG. 2. Predicted amino acid sequences of the V3 region of gp120 of the 94 Ethiopian HIV-1 envelope protein sequences, as compared to a derived Ethiopian consensus sequence. The subtype A sequence is shown first, followed by the main group (62 sequences) and the subcluster (32 sequences). Dashes indicate amino acid identity with the consensus sequence; dots indicate gaps introduced for the purpose of alignment; X, unreadable nucleotides (N); asterisks indicate nucleotide positions with significant differences between the main group and the subcluster. Sequence names have been indicated by codes as described in Fig. 1.

samples collected in a certain year, risk group, or site, giving no information regarding possible independent introductions of these sequences.

Figure 2 presents the amino acid sequence alignment of the Ethiopian samples, subdivided into main group and subcluster. All V3 loops are 35 amino acids long and show the characteristic C-subtype pattern of GPGQT at the apex. The single Ethiopian A-subtype sample shows at the apex a C-like GPGQT motif, instead of the consensus A subtype apex (GPGQA). The Ethiopian C-subtype sequences show a high frequency of methionine at position 306, as compared to other C isolates.[13] This is particularly true for the subcluster (30 of 32 sequences, 94%) as compared to the main group (26 of 62 sequences, 42%). The biological consequences of this mutation remain to be elucidated. On the basis of previous studies the amino acid positions 305 and 319 determine the non-syncytium-inducing (NSI) versus syncytium-inducing (SI) phenotype of HIV-1 strains.[14,15] According to the amino acids predicted at these positions in the sequenced Ethiopian isolates, 93 of them would be NSI and possibly 1 would be SI (B95032). Because the majority of sam-

```
Subcluster

            270                 290                 310                 330                 350
             |                   |                   |                   |                   |
             * *     *       *       *
Consensus:  IVIRSENLTNNAKIIIVQLKEPVEIVCTRPNNNTRKSMRIGPGQTFYATGDIIGDIRQAHCNISEETWNKTLQEVGKKLQEHFPNKTI
C_KS95036   ---K-------------H--N--------------R-----------------------Y---N-S---------A-------X---T
C_AS901290  ---K-----D-V-----H--N--------------------------------------Y--H--K--KE--PR-RE--PK---KQNK
C_TS95007   -I-K-----DHF-K---H-E---K------S-------------------E---------------K--HA--KI----L-QSNKSIR
C_KS95068   -----K---D-V-------Q-------I----------------------A---N----------QK---------R-----------
C_TKS92078  -----------V-----H-QQ-E---------------------------G-------------G-K---------R---D-------
C_TKS92083  ---------D-V-----H-QQ-----------------------------G-------------G-K---------R---D----N--
C_B95042    -I-KAK---D--Q-------K--X----------------------------------------GXE--E------R----L------
C_B9595     T--------K-V-----H--N---------S-------------------E--------------KK--N---D---------L-R--
C_B95086    T----------V--------D---------S-------------------E--------------KN------D-SE---K--L-R--
C_KP95043   T----------V--------D---------S-------------------E--------------QN------D-SE---K--L-R--
C_KAZS9242  ---K-----D-------H-----------------------------------------------QA------G--------------
C_KAZS9258  ---K-------------H-----------------------------------------------QK------------L--------
C_KS95047   ---------D---T---H--------------------------------A--------------K---------------R------
C_B9562     ---K-------V-T---H--D--K------G-------------------AV-------------XD------------HDR--XR--
C_B95032    -----Q---D----------DH-D--------------------------R--------------QK--Q------------Y-----
C_B9545     ---------D---T---H---------S----------------------E----------I---Q----------------Y--T--
C_TP95046   ---K-------------H---------------------------------------R-------T----------E-----------
C_TKS92102  ---K-----D-V-----H------------G------------------------------E---QV-------------------S-
C_KS95020   ---K----XD-V-----H--K----------------I---------------------------QK--E----I-E--K------S-
C_AS901269  ---K-----D-------H--K--------------------------------------------K---E---G-----------R--
C_AS901252  ---K-------------H--K---------------------------------N----Y-----K---E---G--------------
C_TKS92059  ---K-----D-V-----H-EK--------------------------------------------KM--E---G--------------
C_AS90601   ---K-----D-------H--K----------------V---------------------------K---E---R-----------R--
C_2S892733  ---K-----D-V-----H--K-----------------------------G--------Y-----K---E---G------K-------
C_B95077    ---K-----D-V-T---H--N-------X-G---------------------T----------X-K---D---G--E--------X--
C_KAZS9204  ---------D-V-----H--K-------------------------------------------GKA--E---G-----------S--
C_AS90589   ---K-----D-V-----H-EK-------------K---------------G------------------E------E--L--------
C_2S891361  ---K-----D-V-----H--N--G---------K----------------N---------------A------------L------I-
C_B9599     T--------D-V-----H--D--A------------------------------------------K--N----------------S-
C_B9585     -------I---V-T---H--D-----------------------------A------------X-Q---D---X-----E----GR--
C_TKS92021  T----K---D-T-----H--N--------------------------------------------RA--D------E---G-------
C_TS95009   -----------------H--K-----------------------------G-----------V---A-DE------E--HK----N--
```

FIG. 2. *Continued.*

ples collected in this study are from asymptomatic HIV-1-infected individuals, this would be expected.

ACKNOWLEDGMENTS

We thank the National AIDS Control Program Referral Laboratory at the Ethiopian Health and Nutrition Research Institute (NACP, EHNRI) for sera collected before 1995, and the Region 14 Health Bureau and Ethiopian Red Cross Society, National Blood Transfusion Service (ERCS-NBTS), Addis Ababa, for collection of sera in 1995. This work was financially supported by The Netherlands Ministry of Development Cooperation through the Ethiopian–Netherlands AIDS Research Project (ENARP). The Ethiopian–Netherlands AIDS Research Project (ENARP) is a collaborative effort of the Ethiopian Health and Nutrition Research Institute (EHNRI), the Municipal Health Service of Amsterdam, the Department of Human Retrovirology of the Academic Medical Centre (AMC), and the Central Laboratory of the Blood Transfusion Service (CLB), Amsterdam, The Netherlands.

REFERENCES

1. Louwagie J, Janssens W, Mascola J, Heyndrickx L, Hegerich P, van der Groen G, McCutchan FE, and Burke DS: Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. J Virol 1995;69:263–271.
2. Myers G, Korber B, Wain-Hobson S, Jeang K-T, Henderson LE, and Pavlakis GN (eds.): *Human Retroviruses and AIDS*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, 1994.
3. Kostrikis LG, Bagdades E, Cao Y, Zhang L, Dimitriou D, and Ho DD: Genetic analysis of human immunodeficiency virus type 1 strains from patients in Cyprus: Identification of new subtype designated subtype I. J Virol 1995;69:6122–6130.
4. Murphy E, Korber B, Georges-Courbot M, You B, Pinter A, Cook D, Kieny M, Geogres A, Mathiot C, Barre-Sinoussi F, and Girard M: Diversity of V3 region sequences of human immunodeficiency viruses type 1 from the Central African Republic. AIDS Res Hum Retroviruses 1993;9:997–1006.
5. Nkengasong JN, Janssens W, Heyndrickx L, Fransen K, Ndumbe PM, Motte J, Leonaers A, Ngolle M, Ayuk J, Piot P, and van der Groen G: Genetic subtypes of HIV-1 in Cameroon. AIDS 1994; 8:1405–1412.
6. Eshete H and Sahlu T: Review article: The progression of HIV/AIDS in Ethiopia. Ethiop J Health Dev 1996;10:179–190.
7. Zachar V, Goustin AS, Zacharova V, Hager H, Koppelhus U, Womble DD, Liu X, Bambra C, Nyongo A, and Ebbesen P: Genetic polymorphism of envelope V3 region of HIV type 1 subtypes A, C, and D from Nairobi, Kenya. AIDS Res Hum Retroviruses 1996; 12:75–78.
8. Ayehunie S, Johansson B, Sonnerborg A, Zewdie DW, Britton S, and Strannegard O: Sequence analysis of selected regions of the *env* (V3 loop and gp41) and *gag* (p7) reading frames of Ethiopian human immunodeficiency virus type 1 strains. Arch Virol 1993; 128:229–239.
9. Boom R, Sol CJA, Salimans MMM, Jansen CL, Wertheim-van Dillen PME, and van der Noordaa J: A rapid and simple method for purification of nucleic acids. J Clin Microbiol 1990;28:495–503.
10. Higgins DC, Bleasby A, and Fuchs R: CLUSTAL V: Improved software for multiple sequence alignment. Comput Appl Biosci 1992;8:189–191.
11. Saitou N and Nei M: The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–425.
12. Kuiken LC, Cornelissen TEM, Zorgdrager F, Hartman S, Gibbs

JA, and Goudsmit J: Consistent risk group-associated differences in human immunodeficiency virus type 1 vpr, vpu and V3 sequences despite independent evolution. J Gen Virol 1996; 77:783–792.

13. Foley B and Korber B: Global variation in the HIV-1 V3 region. In: *Human Retroviruses and AIDS: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences.* Los Alamos National Laboratory, Los Alamos, New Mexico, 1995.
14. De Wolf F, Hogervorst E, Goudsmit J, Fenyo E-M, Rubsamen-Waigmann H, Holmers H, Galvao-Castro B, Karita E, Wasi C, Sempala SDK, Baan E, Zorgdrager F, Lukashov V, Osmanov S, Kuiken C, Cornelissen M, and the WHO Network for HIV Isolation and Characterization: Syncytium-inducing and nonsyncytium-inducing capacity of human immunodeficiency virus type 1 subtypes other than B: Phenotypic and genotypic characteristics. AIDS Res Hum Retroviruses 1994;10:1387–1400.
15. Fouchier RAM, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, and Schuitemaker H: Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. J Virol 1992; 66:3183–3187.

Address reprint requests to:
Tobias F. Rinke de Wit
Ethiopian–Netherlands AIDS Research Project (ENARP)
Ethiopian Health and Nutrition Research Institute (EHNRI)
P.O. Box 1242
Addis Ababa, Ethiopia