PHYLOGENETIC ANALYSIS OF HUMAN IMMUNODEFICIENCY VIRUS 1 IN GHANA

J. TAKEHISA[1], M. OSEI-KWASI[2], N. K. AYISI[2], O. HISHIDA[1], T. MIURA[1], T. IGARASHI[1], J. BRANDFUL[2], W. AMPOFO[2], V. B. A. NETTY[3], M. MENSAH[4], M. YAMASHITA[1], E. IDO[1], M. HAYAMI[1]*

[1]Laboratory of Pathogenic Virus, Institute for Virus Research, Kyoto University, Sakyo-ku, Kyoto 606, Japan; [2]Noguchi Memorial Institute for Medical Research, University of Ghana, P. O. Box 25, Legon, Accra, Ghana; [3]St. Joseph's Hospital, Koforidua, Ghana; [4]St. Martin's Catholic Hospital, Agomenya, Ghana

Received January 7, 1997; revised January 31, 1997

Summary. – Eleven human immunodeficiency virus 1 (HIV-1) isolates from Ghanaian acquired immunodeficiency syndrome (AIDS) or AIDS-related complex (ARC) patients obtained by our serosurvey in 1986-1994 were genomically analyzed and phylogenetically compared with other known strains. A phylogenetic tree constructed by analyzing the *env* region indicated that heterogeneous HIV-1 strains were circulating in Ghana and the majority of them (9 of 11 isolates) belonged to clade (subtype) A which is now furiously epidemic in Africa. Another isolate (1 of 11) belonged to clade D, and the remaining one (1 of 11) belonged to "clade G". This "clade G" virus grouped by the *env* analysis belonged to clade A by its *pol* sequence, suggesting an A/G intersubtype recombinant. The characteristic sequences in the V3 tip which have not yet been reported were observed in these Ghanaian isolates, which should be taken into account for future vaccine programs.

Key words: human immunodeficiency virus 1 isolates; Ghana; phylogenetic analysis; recombination

The AIDS epidemiology in West Africa has characteristics of a coexistence of both types of HIV, HIV-1 and HIV-2. Our serosurvey carried out in Ghana in 1986 indicated that there were both HIV-1 and HIV-2 infections and that the latter was predominant (Kawamura *et al.*, 1989). However, in our follow-up surveys in 1990 and 1992, HIV-1 infection was more frequently detected among Ghanaian AIDS patients (Hishida *et al.*, 1994). To find a better way to control the disease in West Africa, it is essential to clarify which clades of HIV-1 have been circulating in the area. In this study, we investigated the molecular epidemiology of HIV-1 in Ghana.

*Corresponding author.

Abbreviations: AIDS = acquired immunodeficiency syndrome; ARC = AIDS-related complex; HIV-1 = human immunodeficiency virus 1; IFA = immunofluorescence assay; PBMC = peripheral blood mononuclear cells; PCR = polymerase chain reaction; RT = reverse transcriptase

Study populations were recruited in selected Ghanaian hospitals in Accra, Koforidua, Agomenya, Asin Fosu, Nkawkaw, Tamale and Kumasi in 1986, 1990, 1992 and 1994. The subjects (1986-1992) enrolled in this study were serologically characterized as described (Hishida *et al.*, 1994) and five specimens collected from HIV-1-seropositive patients in 1994 were added. Peripheral blood mononuclear cells (PBMC) taken from seropositive patients were separated by Ficoll gradient sedimentation, and co-cultivated with virus-free human PBMC or a human T-lymphoid cell line, Molt4#8 (Kikukawa *et al.*, 1986) or M8166 (Clapham *et al.*, 1987). The cultures were periodically checked by immunofluorescence assay (IFA) for expression of viral antigens using autologous and reference anti-HIV-1 sera. Antigen-positive cultures were subsequently examined for virus production by measuring reverse transcriptase (RT) activity in the supernatants as previously described (Willey *et al.*, 1988). From the seropositive patients, one HIV-1 strain (GH3) was isolated in 1986, five HIV-1 strains (GH8, 9, 11-13) were isolated in 1992, and another five HIV-1 strains (GH14-18) were isolated in 1994. Chromosomal DNA was extracted from the virus-infected cells using glass milk powder (Prep-A-Gene DNA purification kit; Bio-Rad, Hercules, CA), and a part of the gp120

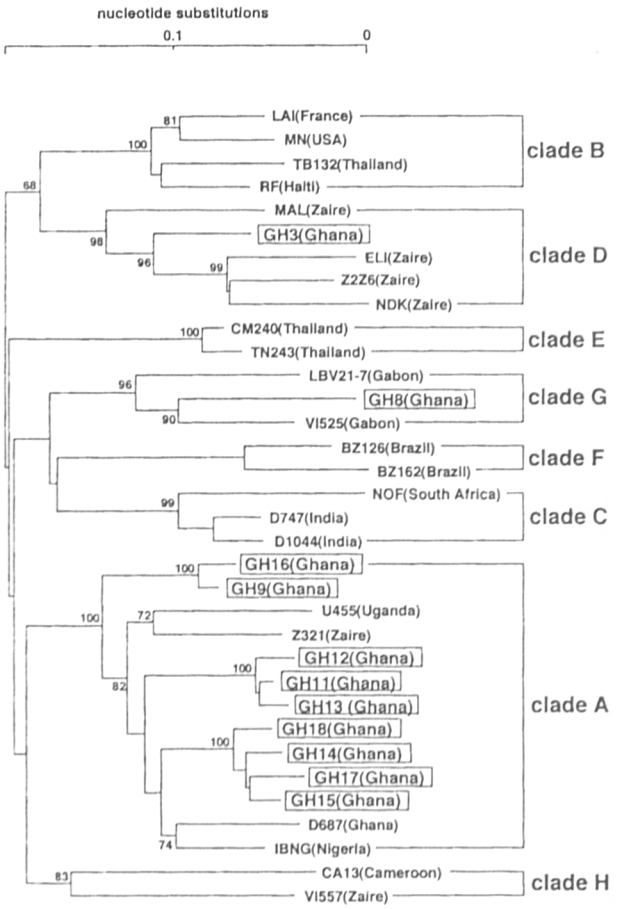

Fig. 1
Phylogenetic tree of various HIV-1 isolates of group M
The tree was constructed from a part of the *env* sequences (approximately 390 bp). Isolate SIVcpzGAB (Gabon) was used as outgroups to root the tree. Clades (subtypes) are indicated by brackets, and the Ghanaian isolates are boxed.

gene of HIV-1 covering one-fourth of the C2, the V3 domain and the whole length of the C3 region corresponding to nt 6613-6999 in HIV-1_{LAI} was amplified by polymerase chain reaction (PCR). Two oligonucleotide primers, F1 (5'-CAACTCAACTGCTGT-TAAATGG-3') and R1 (5'-CCCTCATATCTCCTCCTCCAGG-3') were used. The purified PCR products were subcloned into the *Sma*I site of pUC119 and sequenced by the dideoxy chain termination method using an automated DNA sequencer (373A; Applied Biosystems, Foster City, CA). We verified the mutations, caused by a misincorporation by *Tth* DNA polymerase (Toyobo Co. Ltd., Mishima, Japan), by sequencing at least five clones of the respective PCR products. DNA sequences were aligned by the ODEN program package at the National Institute of Genetics, Tokyo, Japan. Nucleotide substitutions were estimated by the six-parameter method (Gojobori *et al.*, 1982), a phylogenetic tree was constructed by the neighbor-joining method, and its reliability was estimated by 100 bootstrap replications (Felsenstein, 1985). The GenBank accession numbers of respective nucleotide sequences used in this study are U67049 to U67054 and U75457 to U75461.

A phylogenetic tree based on sequences of the central portion of gp120 gene is given in Fig. 1. As shown in the tree, nine (GH9, GH11-18) of eleven HIV-1 isolates belonged to clade A, the geographical distribution of which ranges widely throughout Central and West Africa. Meanwhile, isolate GH3 belonged to clade D, whose distribution is rather limited to Central Africa. The other isolate (GH8) was characteristic because it belonged to clade G, an identified subgroup isolated from Gabon (Janssens *et al.*, 1994).

Fig. 2 shows all of the deduced amino acid sequences from the analyzed nucleotide sequences in the *env* region. For convenient comparison, the results are compiled by each clade. The V3 loops of all Ghanaian HIV-1 isolates [GH9 and GH11-18 (clade A), GH3 (clade D), and GH8 (clade G)] were commonly composed of 35 amino acid residues (cysteine to cysteine). Characteristic tetrameric sequences of the Ghanaian strains were observed at the tip of the V3 loop. In clade A of the African isolates, GPGQ was previously shown to be a predominant tetrameric sequence (Myers *et al.*, 1995). However, the sequences of isolates GH11 and GH13 (clade A) were commonly GLGH, which has not been reported previously, even in other clades. In addition, isolate GH12 (clade A) had a unique tetrameric sequence, GPGH, and isolates GH14, 15, 17 and 18 (clade A) possessed commonly GLGR, both of which were rare in this clade. Isolate GH8 (clade G) had GTGR, which was very rare among all the HIV-1 isolates so far analyzed except isolate VI525 from Gabon (Janssens *et al.*, 1994).

In this study, most of the Ghanaian HIV-1 isolates were grouped into either clade A or D based on the *env* (V3C3) analysis. The same grouping was also obtained based on the *pol* (integrase) analysis (data not shown). However, we found an exceptional case in which isolate GH8 belonged to clade G by its *env* sequence, whereas it was grouped to clade A by its *pol* sequence. One possible explanation for this discordance is that "clade G" viruses are A/G intersubtype recombinants between an "A virus" and a "G virus" (while there has not yet been any identification of an ancestral "G virus"). In fact, some "clade G" viruses were reported to possess mosaic env genes (A/G) (Gao *et al.*, 1996). It is likely that "clade G" viruses are products of multiple recombinational events and therefore they exhibit complex genome structures characterized by multiple cross-overs. As far as the *env* analysis is concerned, "clade G" viruses have also been reported from Gabon, Central African Republic, Uganda, Rwanda, Nigeria, Bénin, Russia, and Taiwan (Myers *et al.*, 1995), suggesting that Central African viruses may be sources of epidemic infections in other areas. Thus, the origin of the "clade G" virus has become an interesting subject which requires further investigation.

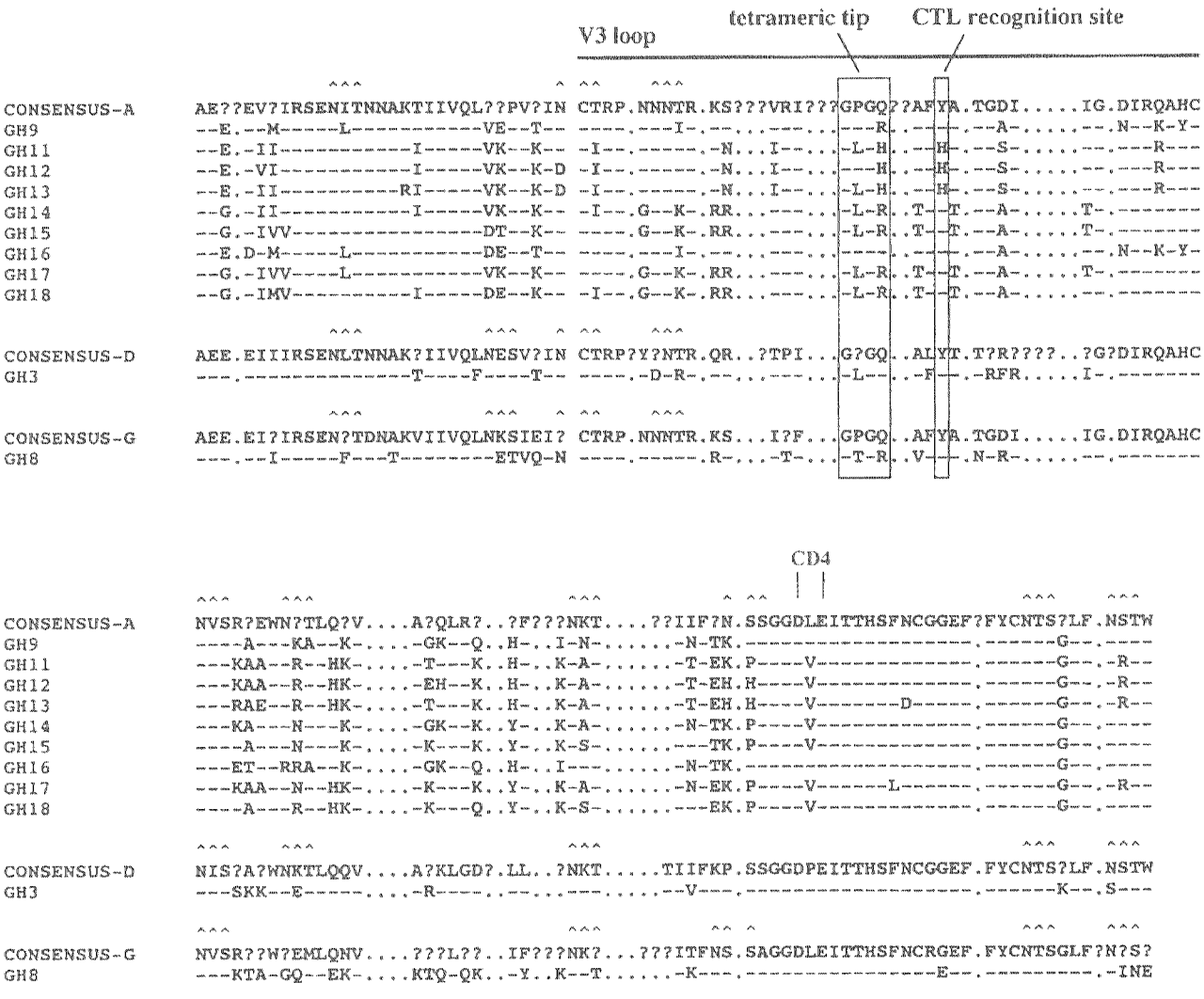

Fig. 2

Alignment of deduced amino acid sequences of the V3 region of the Ghanaian isolates

The consensus amino acid sequences of clades A, D and G based on the Los Alamos database are given on the top line using the single-letter codes. The Ghanaian isolates are denoted by the letters GH. Dashes (-) indicate identities with the consensus sequence. Dots (.) indicate gaps. Signals (^) show positions of potential N-linked glycosylation sites in the consensus sequence. The V3 loop is indicated with a bar.

According to the seroepidemiological survey in Ghana in 1986, the number of cases of HIV-2 infection was almost a 3-fold of that of HIV-1 infection (Hishida *et al.*, 1994). However, the proportion of HIV-1 to HIV-2 in Ghana increased thereafter and HIV-1 infection became predominant in 1990 and 1992. It is likely that HIV-1 was introduced into this country later than HIV-2, and HIV-2 infection was overwhelmed by HIV-1 infection in recent years (Hishida *et al.*, 1994; Ayisi *et al.*, 1995). In fact, the present molecular epidemiological findings make clear that heterogeneous HIV-1 strains have been circulating in Ghana and the majority of them belonged to clade A which is now furiously epidemic in Africa. In addition, the analysis of the HIV-1 *env* region has revealed that Ghanaian isolates possess characteristic sequences in the V3 tip that have not yet been reported, and we have obtained evidence that such genotypes are spreading in Ghana. These molecular data concerning the genetic variability of the envelope glycoprotein of HIV-1 will be very valuable for future vaccine studies in West Africa.

Acknowledgements. The authors are greatly indebted to Drs. Sr. E. Gorges, Catholic Mission Hospital, Battor; Y. Anti, Wenchi Methodist Hospital, Wenchi; K.D. Anane and M. Kyere, St Dominic's Hospital, Akwatia; P. Topor, Catholic Hospital, Asin Fosu; R. DeCiwa and C. Aikins, St Francis Xavier's Hospital, Asin Fosu; F.A.A. Ameyaw and S.B. Ofori, Atua Government Hospital,

Atua; A. Amuzu, Holy Family Hospital, Nkawkaw; M.G. Addy, Doku Hospital, Greater Accra, Ghana, for collecting the samples. We would like to thank Prof. F.K. Nkurumah, Director of the Noguchi Memorial Institute for Medical Research (NMIMR), Accra, Ghana, for supporting this study, and all the staff of the Virology Unit at NMIMR for technical assistance.

References

Ayisi NK, Mensah M, Ishikawa K, Sata T (1995): Serological, virological, and polymerase chain reaction studies of HIV type 1 and HIV type 2 infections in Ghanaian patients with AIDS and AIDS-related complex. *AIDS Res. Hum. Retroviruses* **11**, 319–321.

Clapham PR, Weiss RA, Dalgleish AG, Exley M, Withby D, Hogg N (1987): Human immunodeficiency virus infection of monocytic and T-lymphocytic cells: receptor modulation and differentiation induced by phorbol ester. *Virology* **158**, 44–51.

Felsenstein J (1985): Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Gao F, Morrison SG, Robertson DL, Thornton CL, Craig S, Karlsson G, Sodroski J, Morgado M, Galvao-Castro B, von Briesen H, Beddows S, Weber J, Sharp PM, Shaw GM, Hahn BH, WHO, NIAID Networks for HIV isolation and characterization (1996): Molecular cloning and analysis of functional envelope genes from human immunodeficiency virus type 1 sequence subtypes A through G. *J. Virol.* **70**, 1651–1667.

Gojobori T, Ishii K, Nei M (1982): Estimation of average number of nucleotide substitutions when the rate of substitution varies with nucleotide. *J. Mol. Evol.* **18**, 414–423.

Hishida O, Ayisi NK, Brandful J, Ampofo W, Osei-Kwasi M, Ido E, Igarashi T, Takehisa J, Miura T, Miyazaki A, Hayami M (1994): Serological survey of HIV-1, HIV-2 and human T-cell leukemia virus type 1 for suspected AIDS cases in Ghana. *AIDS* **8**, 1257–1261.

Janssens W, Heyndrickx L, Fransen K, Motte J, Peeters M, Nkengasong JN, Ndumbe PM, Delaporte E, Perret JL, Atende C, Piot P, van der Groen G (1994): Genetic and phylogenetic analysis of *env* subtype G and H in central Africa. *AIDS Res. Hum. Retroviruses* **10**, 877–879.

Kawamura M, Ishikawa K, Mingle JAA, Osei-Kwasi M, Afoakwa SN, Nettey VBA, Chosa T, Hayami M (1989): Immunological reactivities of Ghanaian sera with HIV-1, HIV-2, and simian immunodeficiency virus SIVagm. *AIDS* **3**, 609–611.

Kikukawa R, Koyanagi Y, Harada S, Kobayashi N, Hatanaka M, Yamamoto N (1986): Differential susceptibility to the acquired immunodeficiency syndrome retrovirus in cloned cells of human leukemic T-cell line Molt-4. *J. Virol.* **57**, 1159–1162.

Myers G, Korber B, Hahn BH, Jeang KT, Mellors JW, McCutchan FE, Henderson LE, Pavlakis GN (1995): *Human Retroviruses and AIDS 1995*. Los Alamos National Laboratory, New Mexico.

Willey RL, Smith DH, Laskey LA, Theodore TS, Earl PL, Moss B, Capon DJ, Martin MA (1988): *In vitro* mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity. *J. Virol.* **62**, 139–147.