

The purpose of this study was to investigate the prevalence of subtypes A and C, and the existence of the recombinants of both subtypes in Likasi, located in the southeast of the DRC, 200 km from Lubumbashi in an area bordering Zambia.

Blood samples (10 ml) were collected from 27 HIV-1 infected individuals in February and September 2001. Plasma and buffycoat were separated and stored at -80°C until use. The plasma samples were screened for HIV antibodies with a commercial particle agglutination (PA) test kit (Serodia-HIV, Fujirebio, Tokyo, Japan).

Genomic DNA was extracted from buffycoat of the seroreactive samples using a Qiagen DNA extraction kit (Qiagen, Hilden, Germany). A part of the *pol* gene coding integrase (IN) (corresponding to 4493–4780 nt in HIV-1_{HXB2}) and *env* gene covering C2V3 (corresponding to 6975–7520 nt in HIV-1_{HXB2}) was amplified with nested polymerase chain reaction (PCR) using the primers unipol 5 (5'-TGGGTACCAGCACACAAAGGAATAGGAGGAAA-3')/unipol 6 (5'-CCA-CAGCTGATCTCTGGCCTTCTCTGTAATAGACC-3') and M5 (5'-CCAATTCCCATACATTATTGTGCCCCAGCTGG-3')/M10 (5'-CCAATTGTCCCTCATATCTCCTCCTCCAGG-3'), respectively, in the first round, and unipol 1 (5'-AGTGGATT-CATAGAAAGCAGAAAGT-3')/unipol 2 (5'-CCCCTATTCC-TTCCCTTCTTTAAAA-3'), and M3 (5'-GTCAGCACAG-TACAATGACACATGG-3')/M8 (5'-TCCTTGGATGGGAG-GGGCATACATTGC-3'), respectively, in the second round.¹¹ Nested PCR was performed with an AmpliTaq Gold PCR kit (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Amplification was done with one cycle at 95°C for 10 min, and 35 cycles at 95°C for 30 sec, 45°C for 30 sec (for the *pol* region) or 55°C for 30 sec (for the *env* region), and 72°C for 1 min, with a final extension of 72°C for 10 min. Samples that could not be amplified with the *env*-C2V3 primers were analyzed with *env*-gp41-specific primers.¹² PCR amplification was confirmed by visualization with ethidium bromide staining of the gel. The PCR products were cloned by using a TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) and plasmid DNA for sequencing was prepared by a plasmid miniprep kit (Sigma, Hilden, Germany). DNA sequencing was carried out using dye-deoxy terminator chemistry on an ABI 310 automatic sequencer (Applied Biosystems, Foster City, CA). We sequenced at least 12 clones to obtain a consensus sequence. Sample DNA sequences were aligned with subtype reference sequences from the Los Alamos database by CLUSTAL W (version 1.81) with subsequent inspection and manual modification. The frequency of nucleotide substitution in each base of the sequences was estimated by the Kimura two-parameter method.¹³ Phylogenetic trees were constructed by the neighbor-joining method, and its reliability was estimated by 1000 bootstrap replications.¹⁴ All alignments were gap stripped for the generation of trees. The profile of the tree was visualized by Treeview version 1.6.5.

The phylogenetic trees based on *env*-C2V3 and *pol*-integrase sequences were constructed with representative HIV-1 strains of each subtype and CRF as a reference (Fig. 1). Out of the 27 samples from Likasi; 24 (88.9%) were found to be positive for HIV-1 PCR with *pol*-IN and *env*-C2V3 primers. The remaining 3 (11.1%) samples were negative for PCR with the *env*-C2V3 primers used in this study. The phylogenetic tree based on *env*-C2V3 sequences (Fig. 1A) showed that out of the 24

samples 13 were subtype A, three C, one D, two G, one CRF01_AE, and four U (unclassified). The outcome of the phylogenetic analysis of the *pol*-IN gene is shown in Figure 1B, and revealed that 10 were subtype A, three C, one D, four G, three H, three J, two CRF02_AG, and one U. Thus, 15 (62.5%) had a discordant subtype or CRF designation: one subtype A/G (*pol/env*), four A/U, three G/A, one G/CRF01, three H/A, one J/C, and CRF02/A, and one U/A (Tables 1 and 2). Nine (37.5%) strains had a concordant subtype or CRF designation: one subtype A, two C, one D, and one CRF02/G (Tables 1 and 2). In 14 of these 15 strains, subtype A was involved in recombination events, and among the strains with a discordant subtype or CRF designations A/U (*pol/env*) was by far the most common, followed by H/A and G/A recombination. Additional PCR analyses on three samples with a *pol/env* (+/-) profile were carried out with groups M, N, and O primers for *env*-gp41. Phylogenetic analysis based on *env*-gp41 sequences revealed the presence of two profiles: two J/J (*pol*-IN/*env*-gp41) and one C/G (Tables 1 and 2).

In the current study we found that a high proportion (16/27; 59.3%) of HIV-1 strains in Likasi were intersubtype recombinants. This is higher than that reported in other regions of the country (29–44%).^{5,8,15} In Lubumbashi, a city on the southeast border of the DRC with Zambia, subtype C was reported to be predominant (51.9%), followed by subtype A (22.1%).¹⁵ Our data highlight a high prevalence of subtype A (37.0% and 52.5%) and low prevalence of subtype C (7.4% and 12.5%) in Likasi based on *env*-C2V3 and *pol*-IN analyses, respectively. The persisting civil war and population displacement from the east on the border with Ruanda and Burundi to the southern area close to Zambia could be linked to the change of HIV-1 distribution in the southeast of the DRC.¹¹

Two HIV-1 strains from Likasi (00CD009 and 01CD208) significantly clustered with CRF02_AG reference strains (with 97.4% bootstrap value) (Fig. 1B). This is the first report of CRF02_AG in the DRC, suggesting that CRF02_AG is spreading into Central Africa. CRF02_AG and subtype A represent 70–80% of circulating HIV-1 strains in West and West-Central Africa.² However, the epidemiological significance of CRF02_AG in the DRC is yet to be investigated. The high proportion of unclassified strains (16.6%) and intersubtypic recombinants (52.3%) among HIV-1 strains circulating in the DRC indicates an old and mature epidemic of HIV-1 in the DRC.

SEQUENCE DATA

The nucleotide sequences in this study were submitted to GenBank and are available under the following accession numbers: *pol*-IN (288 bp), AY661750–AY661776; *env*-C2V3 (approximately 550 bp), AY675589–AY675612; and *env*-gp41 (405 bp), AY673112–AY673114.

ACKNOWLEDGMENTS

This work was supported in part by International Scientific Research Program Grants 12376002 and 14256005 from Monbukagakusho (Ministry of Education and Science).

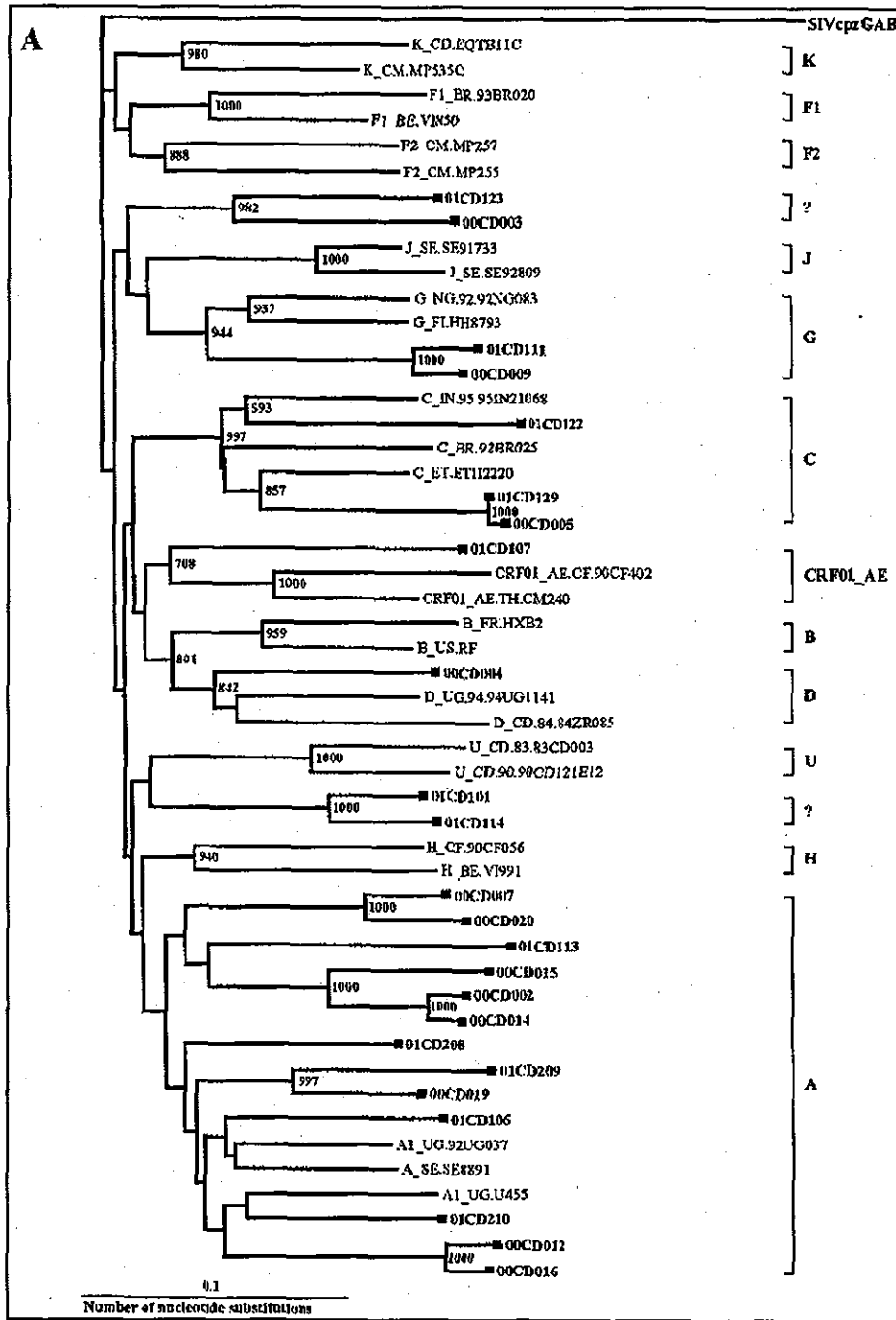


FIG. 1. Phylogenetic trees based on a part of the *env*-C2V3 gene (approximately 550 bp) (A) and *pol*-IN gene (288 bp) (B) of 24 HIV-1 strains from the southeastern Democratic Republic of Congo with reference sequences of representative subtypes/CRF. The bootstrap value at each node represents the number among 1000 bootstrap replicates that supports the branching order. Bootstrap values of 70% or higher are shown. Brackets on the right represent the major group *M* subtypes. Newly derived sequences (shown in bold) are marked with a filled square (■).

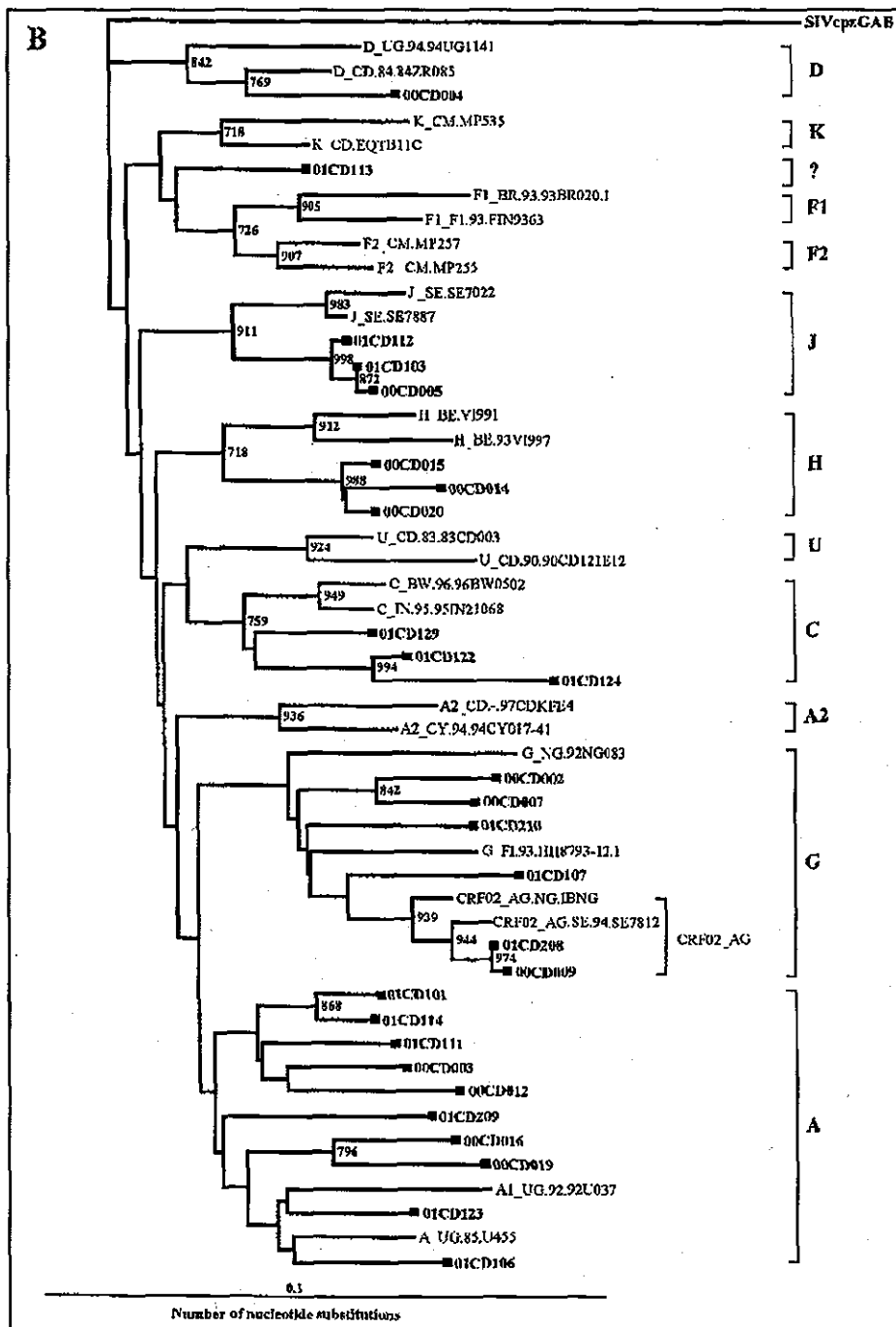


FIG. 1. Continued

TABLE 1. SUBTYPES IN *pol*-INTEGRASE AND *env*-(C2V3/ OR -gp41) FOR 27 HIV-1 STRAINS

ID number DRC	<i>pol-IN</i> (288 bp)	<i>env-C2V3</i> (550 bp)	<i>env-gp41</i> (405 bp)
00CD002	Subtype G	Subtype A	
00CD003	A	U	
00CD004	D	D	
00CD005	J	C	
00CD007	G	A	
00CD009	CRF02_AG (G)	G	
00CD012	A	A	
00CD014	H	A	
00CD015	H	A	
00CD016	A	A	
00CD019	A	A	
00CD020	H	A	
01CD101	A	A	
01CD103	J	ND ^a	J
01CD106	A	A	
01CD107	G	CRF01_AE (E)	
01CD111	A	G	
01CD112	J	ND	J
01CD113	U	A	
01CD114	A	A	
01CD122	C	C	
01CD123	A	U	
01CD124	C	ND	G
01CD129	C	C	
01CD208	CRF02_AG (G)	A	
01CD209	A	A	
01CD210	G	A	

^aND, not detected.TABLE 2. SUBTYPES IN *pol*-INTEGRASE AND *env*-(C2V3) AND/OR -(gp41) GENE FOR 27 HIV-1 STRAINS

Genotypes	<i>pol-IN</i>	<i>env-C2V3</i>	n	Total (%)
Concordant	A	A	5	11 (40.7)
	C	C	2	
	D	D	1	
	J	J ^a	2	
	CRF02 (G)	G	1	
Discordant	A	U	4	16 (59.3)
	A	G	1	
	C	G ^a	1	
	G	A	3	
	G	CRF01 (E)	1	
	H	A	3	
	J	C	1	
	CRF02 (G)	A	1	
U	A	1		

^a*env-gp41* (approximately 405 bp).

REFERENCES

- Robertson DL, Anderson JP, Bradac JA, *et al.*: HIV-1 nomenclature proposal. *Science* 2000;288:55-56.
- Peeters M, Toure-Kane C, and Nkengasong JN: Genetic diversity of HIV in Africa: Impact on diagnosis, treatment, vaccine development and trials. *AIDS* 2000;17:2547-2560.
- Papathanasopoulos MA, Hunt GM, and Tiemessen CT: Evolution and diversity of HIV-1 in Africa—a review. *Virus Genes* 2003;26:151-163.
- Janssens W, Buve A, Nkengasong JN, *et al.*: The puzzle of HIV-1 subtypes in Africa [editorial]. *AIDS* 1997;11:705-712.
- Vidal N, Peeters M, Mulanga-Kabeya C, *et al.*: Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *J Virol* 2000;74:10498-10507.
- Mokili JL, Wade CM, Burns SM, *et al.*: Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural Democratic Republic of Congo. *AIDS Res Hum Retroviruses* 1999;15:655-664.
- Mokili JL, Rogers M, Carr JK, *et al.*: Identification of a novel clade of human immunodeficiency virus type 1 in Democratic Republic of Congo. *AIDS Res Hum Retroviruses* 2002;18:817-823.
- Vidal N, Koyalta D, Richard V, *et al.*: High genetic diversity of HIV-1 strains in Chad, West Central Africa. *J Acquir Immune Defic Syndr* 2003;33:239-246.
- Trask SA, Derdeyn CA, Fideli U, *et al.*: Molecular epidemiology of human immunodeficiency virus type 1 transmission in a het-

- crosexual cohort of discordant couples in Zambia. *J Virol* 2002; 76:397-405.
10. Handema R, Terunuma H, Kasolo F, *et al.*: Emergence of new HIV-1 subtypes other than subtype C among antenatal women in Lusaka, Zambia. *AIDS Res Hum Retroviruses* 2001;17:759-763.
 11. Mulanga C, Bazepeo E, Mwamba J, *et al.*: Political and socioeconomic instability: How does it affect HIV? A case study in the Democratic Republic of Congo. *AIDS* 2004;18:833-834.
 12. Yang C, Pieniazek D, Owen SM, Fridlung C, *et al.*: Detection of phylogenetically diverse human immunodeficiency virus type 1 groups M and O from plasma by using highly sensitive and specific generic primers. *J Clin Microbiol* 1999;37:2581-2586.
 13. Kimura M: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotides sequences. *J Mol Evol* 1980;16:111-120.
 14. Saitou N and Nei M: The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;4:406-425.
 15. Vidal N, Mulanga C, Bazepeo E, *et al.*: A constant and extremely high genetic diversity of HIV-1 group M strains in the Democratic Republic of Congo, between 1997-2002 (Abstract No. 246: S243). Proceedings of the 2nd International AIDS Conference on HIV Pathogenesis and Treatment, Paris, 13-16 July 2003.

Address reprint requests to:

Hiroshi Ichimura
Department of Viral Infection and International Health
Graduate School of Medical Science
Kanazawa University
13-1 Takara-machi
Kanazawa 920-8640, Japan

E-mail: ichimura@med.kanazawa-u.ac.jp

Genetic Diversity of HIV Type 1 in Rural Eastern Cameroon

Nicaise Ndembi,*† Jun Takehisa,* Leopold Zekeng,† Eiji Kobayashi,* Charlotte Ngansop,†
Elijah M. Songok,* Seiji Kageyama,* Taichiro Takemura,§ Eiji Ido,§ Masanori Hayami,§
Lazare Kaptue,† and Hiroshi Ichimura, MD, PhD*

Summary: To monitor the presence of genotypic HIV-1 variants circulating in eastern Cameroon, blood samples from 57 HIV-1-infected individuals attending 3 local health centers in the bordering rural villages with Central African Republic (CAR) were collected and analyzed phylogenetically. Out of the 40 HIV-1 strains with positive polymerase chain reaction (PCR) profile for both *gag* and *env*-C2V3, 12 (30.0%) had discordant subtype or CRF designation: 2 subtype B/A (*gag/env*), 1 B/CRF01, 2 B/CRF02, 1 CRF01/CRF01.A, 2 CRF11/CRF01, 1 CRF13/A, 1 CRF13/CRF01, 1 CRF13/CRF11, and 1 G/U (unclassified). Twenty-eight strains (70.0%) had concordant subtypes or CRF designation between *gag* and *env*: 27 subtype A and 1 F2. Out of the remaining 17 HIV-1 strains negative for PCR with the *env*-C2V3 primers used, 10 (58.8%) had discordant subtype or CRF, and 7 (41.2%) had concordant one based on *gag/pol/env-gp41* analysis. Altogether, a high proportion (22/57, 38.6%) of the isolates were found to be recombinant strains. In addition, an emergence of new forms of HIV-1 strains, such as subtype B/A (*gag/env*), B/CRF01 and B/CRF02, was identified. The epidemiologic pattern of HIV-1 in eastern Cameroon, relatively low and high prevalence of CRF02 and CRF11, respectively, was more closely related to those of CAR and Chad than that of other regions of Cameroon, where CRF02 is the most predominant HIV-1 strain. These findings strongly suggest that this part of Cameroon is a potential hotspot of HIV-1 recombination, with a likelihood of an active generation of new forms of HIV-1 variants, though epidemiologic significance of new HIV-1 forms is unknown.

Key Words: HIV, subtypes, CRF, recombinant, hotspot

(*J Acquir Immune Defic Syndr* 2004;00:000-000)

Recent studies have suggested that simian immunodeficiency virus of chimpanzee (SIVcpz) has been introduced into humans at least 3 times, resulting in HIV-1 groups M, N, and O.¹⁻⁴ The cross-species transmission of SIVcpz is now thought to have occurred through humans being exposed to the blood of chimpanzees infected with SIVcpz during hunting and butchering of these chimpanzees in rural areas of the equatorial rain forest of Central Africa.⁴⁻⁶

The epidemiologic pattern of HIV-1 infection in Cameroon, a west-central African country, is characterized by an extensive genetic variability of the virus in terms of co-circulation of all representative major groups and subtypes,¹⁻⁴ though CRF02_AG and subtype A represent 70-80% of HIV-1 infection in the country.⁷⁻¹⁸ HIV-1 group O seems to be endemic in Cameroon, where the frequency is estimated to be 2-5% among HIV-1-infected individuals.^{19,20} Group N, the least spreading virus group among HIV-1 isolates, is represented by only a handful of viruses identified in some villages in Cameroon. Overall seroprevalence of HIV-1 in rural villages has been reported to be relatively low (0.7-6%) compared with urban and semiurban areas (11%) in Cameroon.^{5,6,21} However, most studies have been focused on patients living in major urban areas, and less information is available on the distribution of group M strains circulating in rural areas,⁵ where >60% of Cameroonians reside.

In the east of Cameroon, Chad, relatively low prevalence of CRF02_AG (13.1%) and relatively high prevalence of CRF11_cpx (13.1-18.0%) and subtype D (18.7%) compared with those in neighboring countries, have been reported. In the southeast of Cameroon, the Central African Republic (CAR), a high prevalence of subtype A (69.7%) and CRF01_AE (21.8%) has been reported,^{22,23} showing that the subtype distribution is heterogeneous in west-central Africa.^{24,25}

The purpose of this study was to investigate the genetic diversity of HIV-1 strains circulating in rural and border villages of eastern Cameroon, to compare the result with those of Chad and CAR, and to determine the influence of migration of

Received for publication December 29, 2003; accepted for publication April 5, 2004.

From the *Department of Viral Infection and International Health, Graduate School of Medical Science, Kanazawa University, Kanazawa, Japan; †Laboratoire de Santé Hygiene Mobile, Ministère de la Santé, Yaoundé, Cameroon; ‡Laboratory of Hematology and Virology, Faculty of Medicine and Biomedical Sciences, University of Yaoundé-I, Cameroon; and §Laboratory of Viral Pathogenesis, Institute for Virus Research, Kyoto University, Kyoto, Japan.

Supported in part by International Scientific Research Program grants 12376002 and 14256005 from Monbu-kagakusho (Ministry of Education and Science in Japan).

Reprints: Hiroshi Ichimura, Department of Viral Infection and International Health, Graduate School of Medical Science, Kanazawa University, 13-1 Takara-machi, Kanazawa 920-8640, Japan (e-mail: ichimura@med.kanazawa-u.ac.jp).

Copyright © 2004 by Lippincott Williams & Wilkins

J Acquir Immune Defic Syndr • Volume 00, Number 0, Month 0 2004

TABLE 1. Characteristics and Genetic Data for HIV-1-Infected Subjects Studied in Eastern Cameroon

Sample ID*	Age (y)	Sex	Nationality	Clinical Status	Genetic Subtype†			
					<i>gag</i>	<i>pol</i>	<i>env</i> C2V3	gp41
01CM2213	25	M	CMR	ARC	CRF_01.AE	NA	CRF_01.AEA	NA
01CF2214	30	M	CMR	ARC	G	U	U	NA
01CM2215	34	F	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2216	36	M	CMR	ARC	A	NA	A	NA
01CM2217	26	M	CMR	ARC	CRF_11.cpx	NA	CRF_11.cpx	NA
01CM2218	25	F	CMR	AC	CRF_11.cpx	CRF_11.cpx	ND	U
01CM2219	32	F	CMR	ARC	CRF_11.cpx	NA	CRF_02.AG	NA
01CM2220	29	M	CMR	ARC	CRF_02.AG	NA	A	NA
01CM2222	46	M	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2223	25	M	CMR	ARC	CRF_01.AE	NA	CRF_02.AG	NA
01CM2224	29	F	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2225	31	M	CMR	ARC	B	NA	A	NA
01CM2226	32	M	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2227	30	F	CMR	AC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2228	42	F	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2229	30	F	CMR	AIDS	CRF_11.cpx	NA	CRF_11.cpx	NA
01CM2230	48	M	CMR	ARC	A	NA	A	NA
01CM2231	42	F	CMR	ARC	CRF_02.AG	NA	A	NA
01CM2232	19	F	CMR	AC	B	U	A	U
01CM2234	28	M	CMR	ARC	CRF_11.cpx	NA	CRF_02.AG	NA
01CM2235	33	M	CMR	ARC	B	U	ND	U
01CM2236	44	F	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2237	38	M	CMR	ARC	F2	NA	F2	NA
01CM2238	29	F	CMR	ARC	CRF_13.cpx	NA	CRF_01.AE	NA
01CM2239	25	M	CMR	ARC	CRF_13.cpx	NA	CRF_11.cpx	NA
01CF2240	50	M	CMR	ARC	CRF_02.AG	NA	CRF_13.cpx	NA
01CM2241	30	M	CMR	ARC	CRF_01.AE	CRF_11.cpx	ND	U
01CM2242	48	M	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2243	30	F	CMR	AC	CRF_11.cpx	CRF_11.cpx	ND	CRF_11.cpx
01CM2244	2	F	CMR	ARC	CRF_01.AE	NA	CRF_11.cpx	NA
01CM2246	17	F	CMR	ARC	B	NA	CRF_01.AE	NA
01CM2260	29	F	CMR	ARC	CRF_13.cpx	U	A	CRF_13.cpx
01CM2262	27	F	CMR	ARC	B	NA	CRF_02.AG	NA
01CF2268	23	F	CMR	AC	CRF_02.AG	CRF_02.AG	ND	CRF_02.AG
01CM2269	21	F	CMR	ARC	CRF_11.cpx	CRF_11.cpx	ND	CRF_11.cpx
01CM2270	26	M	CMR	ARC	CRF_02.AG	CRF_02.AG	ND	U
01CM2271	26	F	CMR	ARC	CRF_11.cpx	CRF_02.AG	ND	CRF_11.cpx
01CM2272	27	F	CMR	ARC	CRF_11.cpx	NA	CRF_11.cpx	NA
01CM2273	54	M	CMR	ARC	CRF_11.cpx	NA	CRF_11.cpx	NA
01CM2274	35	F	CMR	ARC	CRF_02.AG	NA	CRF_02.AG	NA
01CM2275	35	M	CMR	ARC	CRF09_cpx	CRF_02.AG	ND	CRF09_cpx
01CM2276	30	F	CMR	ARC	CRF_11.cpx	NA	CRF_11.cpx	NA
01CM2277	28	F	CMR	ARC	CRF_11.cpx	CRF_11.cpx	ND	CRF_11.cpx
01CM2278	37	F	CMR	ARC	B	NA	CRF_02.AG	NA
01CM2280	26	M	CMR	AC	CRF_11.cpx	CRF_02.AG	ND	CRF_02.AG
01CM2281	20	F	CMR	ARC	CRF_02.AG	CRF_02.AG	ND	CRF_02.AG
01CM2284	25	M	CMR	AC	CRF_11.cpx	CRF_11.cpx	ND	CRF_11.cpx
01CF2287	33	F	CMR	AC	CRF_11.cpx	NA	CRF_01.AE	NA
02CM319	18	F	CMR	AC	ND	"O"‡	ND	"O"

*Subject identification number is preceded by year of sample collection; †genotyping of a part of *gag*-p24 (460 bp), *env*-C2V3 domain (approximately 550 bp), *pol*-In (288 bp), and *env*-gp41 (405 bp); ‡Group O from rural center province of Cameroon.

AC, asymptomatic carrier; ARC, AIDS-related complex; CMR, Cameroon; NA, not available; ND, not detected.

people across the border on the genetic diversity of HIV-1 in the eastern part of Cameroon.

MATERIALS AND METHODS

Study Population

The subjects enrolled in this study were HIV-1-infected individuals recruited in September 2001 in 3 semirural areas of

eastern Cameroon near Bertoua, a town on the road connecting the capital cities of Cameroon (Yaounde) and CAR (Bangui). They were 52 Cameroonians and 5 CAR citizens attending 3 local health centers and suspected of or having sexually transmitted infection, tuberculosis, or AIDS. After obtaining informed consent and ethical clearance, 10 mL of blood samples were collected from the 57 patients (27 men and 30 women,

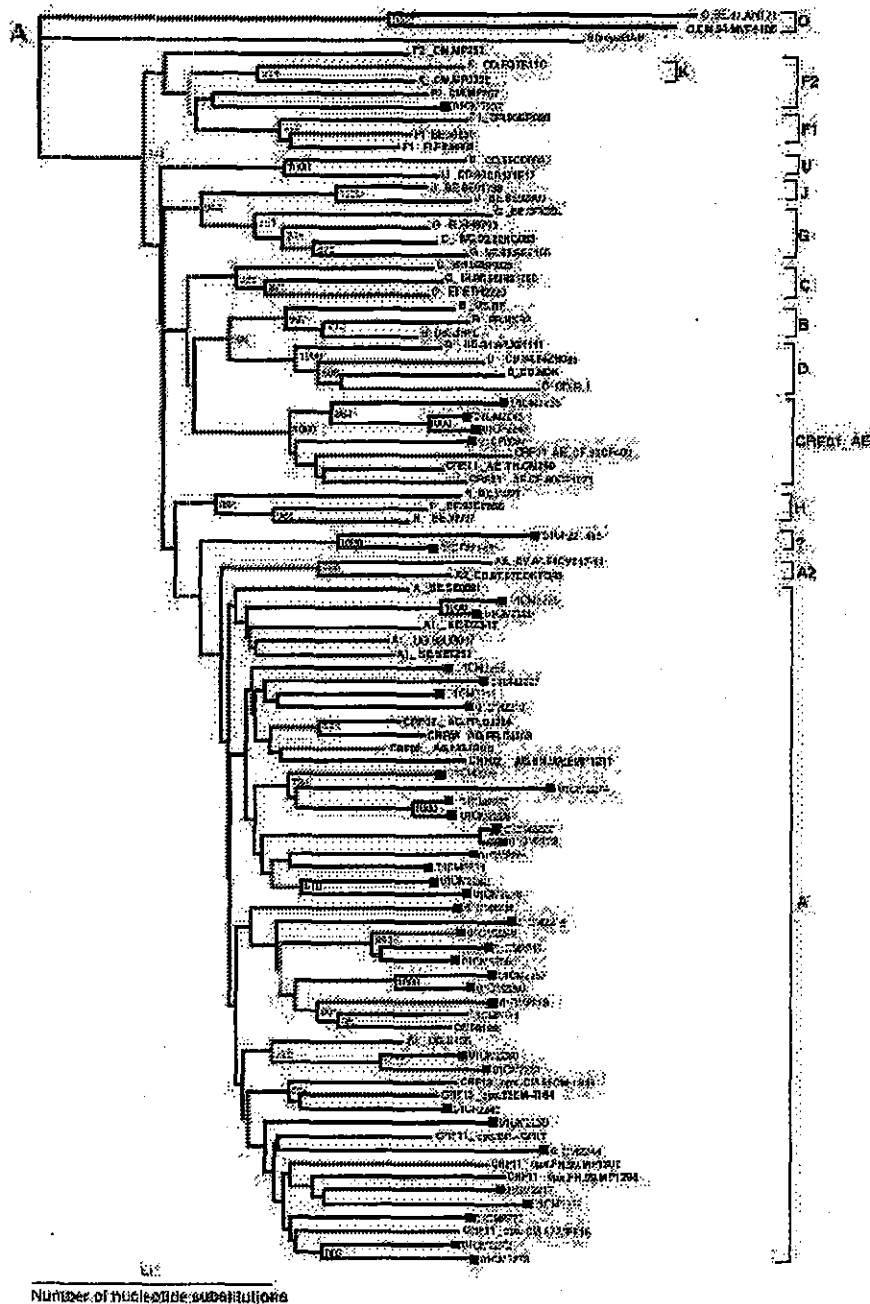


FIGURE 1. Phylogenetic trees based on a part of *env*-C2V3 gene (approximately 550 bp) of 40 HIV-1 strains (A), *gag* gene (approximately 460 bp) of 56 HIV-1 strains (B), *pol*-IN gene (288 bp) of 20 HIV-1 strains (C), and *env*-gp41 gene (approximately 405 bp) of 19 HIV-1 strains (D) from eastern Cameroon with reference sequences of representative subtypes/CRF. The reference sequences used in *gag*, *pol*, and *env* trees are shown in bold. The bootstrap value at each node represents the number among 1000 bootstrap replicates that support the branching order. Bootstrap value of $\geq 70\%$ or higher is shown. Brackets on the right represent the major group M subtypes. Newly derived sequences are marked with a filled square (■). Further analysis of the 34 probable subtype A sample sequences without the rest of the sample sequences is shown in part E.

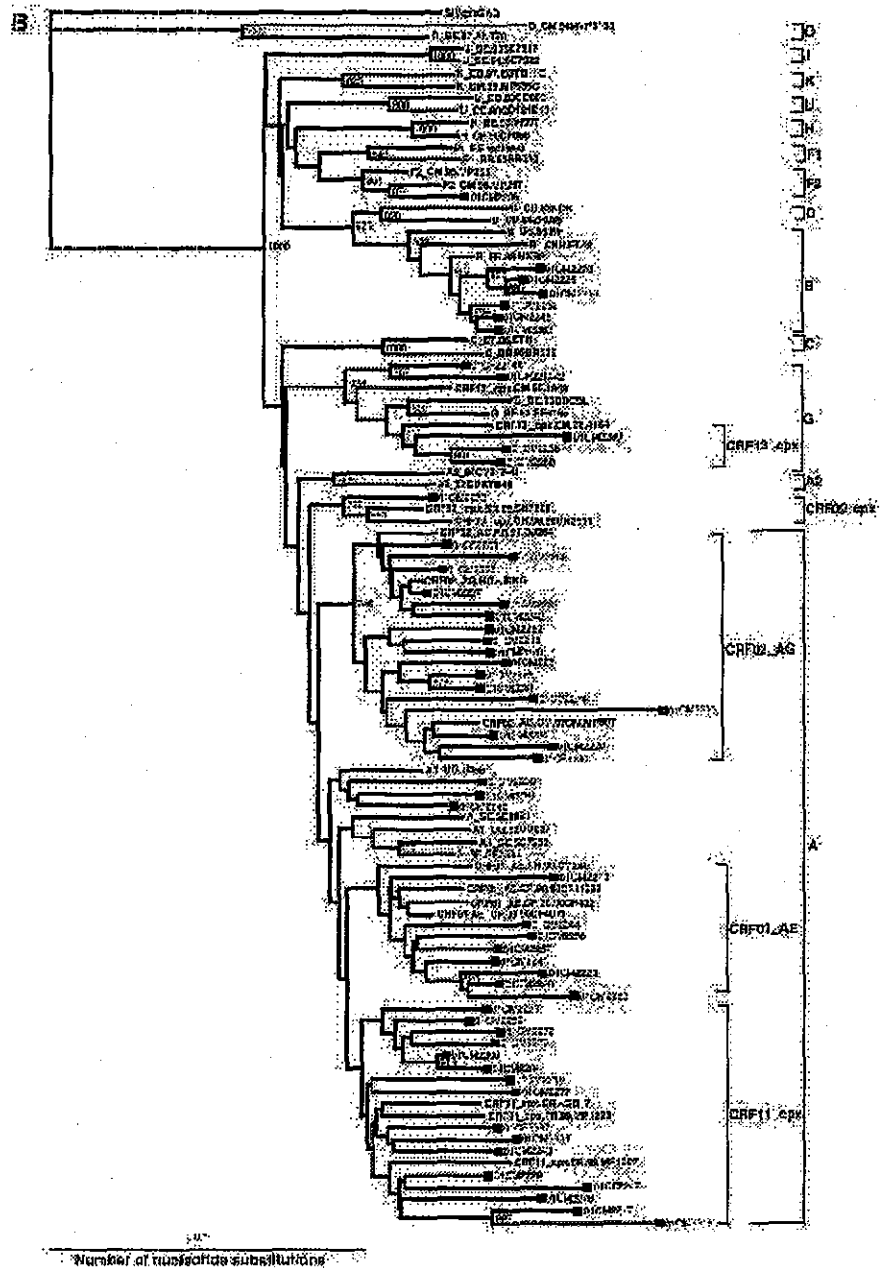


FIGURE 1. (continued) Phylogenetic trees based on a part of *env*-C2V3 gene (approximately 550 bp) of 40 HIV-1 strains (A), *gag* gene (approximately 460 bp) of 36 HIV-1 strains (B), *pol*-1N gene (288 bp) of 20 HIV-1 strains (C), and *env*-gp41 gene (approximately 405 bp) of 19 HIV-1 strains (D) from eastern Cameroon with reference sequences of representative subtypes/CRF. The reference sequences used in *gag*, *pol*, and *env* trees are shown in bold. The bootstrap value at each node represents the number among 1000 bootstrap replicates that support the branching order. Bootstrap value of $\geq 70\%$ or higher is shown. Brackets on the right represent the major group M subtypes. Newly derived sequences are marked with a filled square (■). Further analysis of the 34 probable subtype A sample sequences without the rest of the sample sequences is shown in part E.

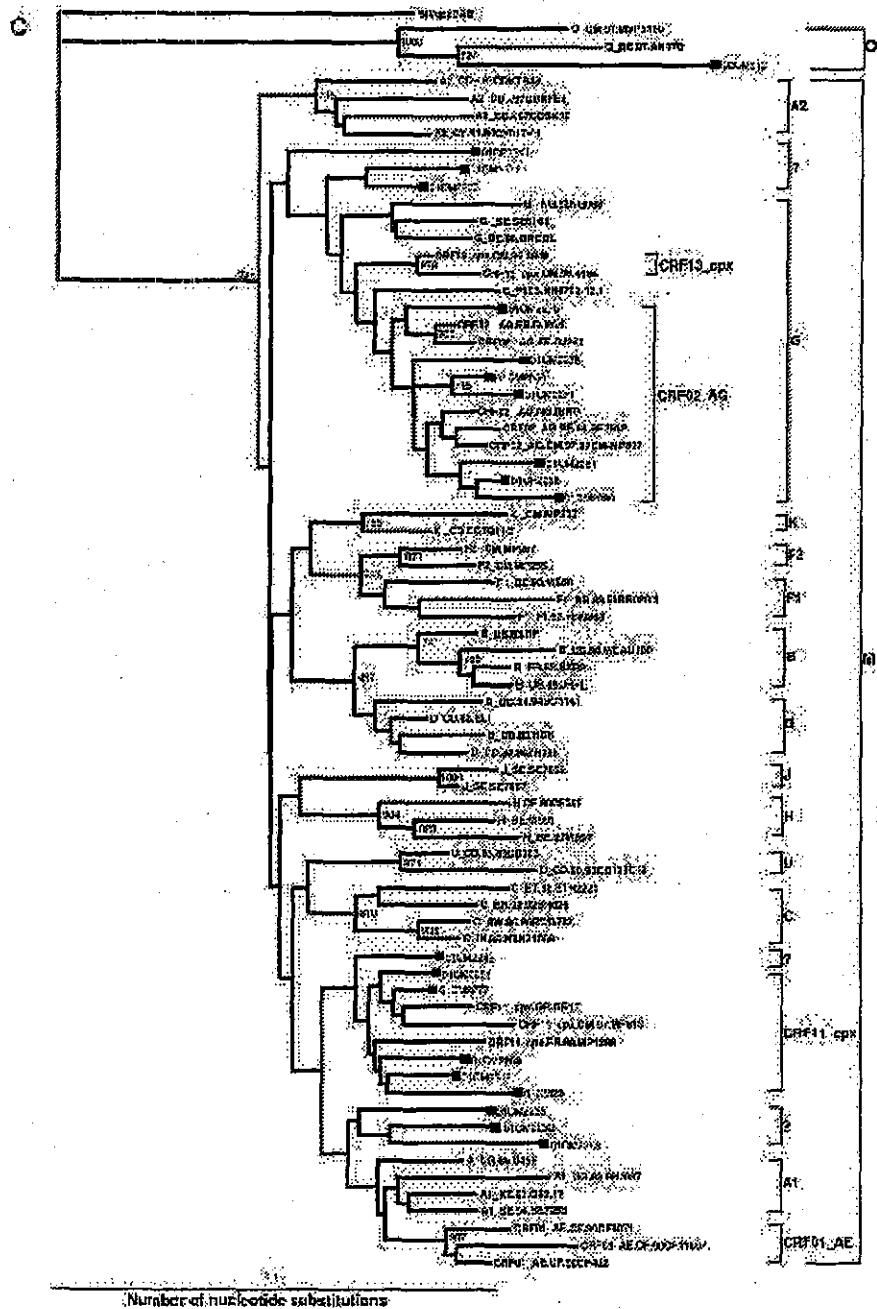
□ mean age \pm SD: 31.5 \pm 10.0 years) (Table 1) and screened for HIV antibodies. The plasma samples were confirmed to be positive for HIV antibodies with a microparticle enzyme immunoassay kit (AxSYM HIV1/2; Abbott Japan) and an immunochromatography assay kit (Determine HIV1/2; Abbott Japan). The peripheral blood mononuclear cells were prepared by Ficoll-Paque Plus (Amersham Biosciences, Uppsala, Sweden) density gradient centrifugation. Genomic DNA was

extracted from peripheral blood mononuclear cells of sero-reactive samples using a DNA extraction kit (Qiagen, Hilden, Germany).

Polymerase Chain Reaction, Cloning, and Sequencing

A part of HIV-1 group M *env* gene covering C2V3 (corresponding to 6975–7520 nt in HIV-1_{HXB2}) was amplified by

FIGURE 1. (continued) Phylogenetic trees based on a part of *env*-C2V3 gene (approximately 550 bp) of 40 HIV-1 strains (A), *gag* gene (approximately 460 bp) of 56 HIV-1 strains (B), *pol*-IN gene (288 bp) of 20 HIV-1 strains (C), and *env*-gp41 gene (approximately 405 bp) of 19 HIV-1 strains (D) from eastern Cameroon with reference sequences of representative subtypes/CRF. The reference sequences used in *gag*, *pol*, and *env* trees are shown in bold. The bootstrap value at each node represents the number among 1000 bootstrap replicates that support the branching order. Bootstrap value of $\geq 70\%$ or higher is shown. Brackets on the right represent the major group M subtypes. Newly derived sequences are marked with a filled square (■). Further analysis of the 34 probable subtype A sample sequences without the rest of the sample sequences is shown in part E.



nested polymerase chain reaction (PCR) with primers M5 (5'-CCAATTCCCATACATTATTGTGCCCCAGCTGG-3') and M10 (5'-CCAATTGTCCTCATATCTCCTCCTC-CAGG-3') in the first round, and M3 (5'-GTCAGCACAG-TACAATGCACACATGG-3') and M8 (5'-TCCTTG-GATGGGAGGGGCATACATTGC-3') in the second round.⁶ A part of *gag* encoding p24 (corresponding to 1596–2016 nt in

HIV-1_{IND2}) was amplified by nested PCR using primers H1G777 (5'-TCACCTAGAACTTTGAATGCATGGG-3') and HIP202 (5'-CTAATACTGTATCATCTGCTCCTGT-3') in the first round, and H1gag1584 (5'-AAAGATG-GATAATCCIGGG-3') and g17 (5'-TCCACATTTCCAA-CAGCCCTTTT-3') in the second round.⁷ A part of *env* gene encoding gp41 (corresponding to 7880–8280 nt in HIV-

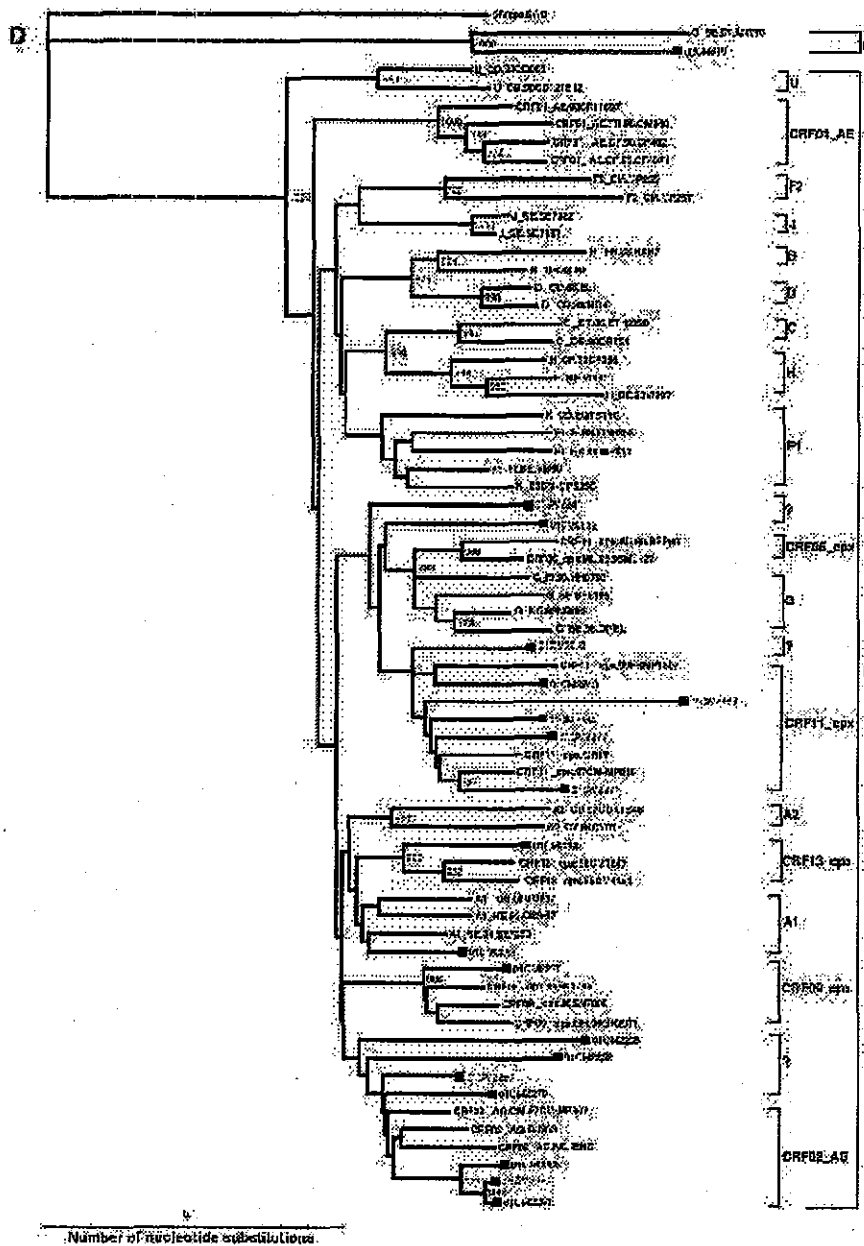


FIGURE 1. (continued) Phylogenetic trees based on a part of *env*-C2V3 gene (approximately 550 bp) of 40 HIV-1 strains (A), *gag* gene (approximately 460 bp) of 56 HIV-1 strains (B), *pol*-IN gene (288 bp) of 20 HIV-1 strains (C), and *env*-gp41 gene (approximately 405 bp) of 19 HIV-1 strains (D) from eastern Cameroon with reference sequences of representative subtypes/CRF. The reference sequences used in *gag*, *pol*, and *env* trees are shown in bold. The bootstrap value at each node represents the number among 1000 bootstrap replicates that support the branching order. Bootstrap value of $\geq 70\%$ or higher is shown. Brackets on the right represent the major group M subtypes. Newly derived sequences are marked with a filled square (■). Further analysis of the 34 probable subtype A sample sequences without the rest of the sample sequences is shown in part E.

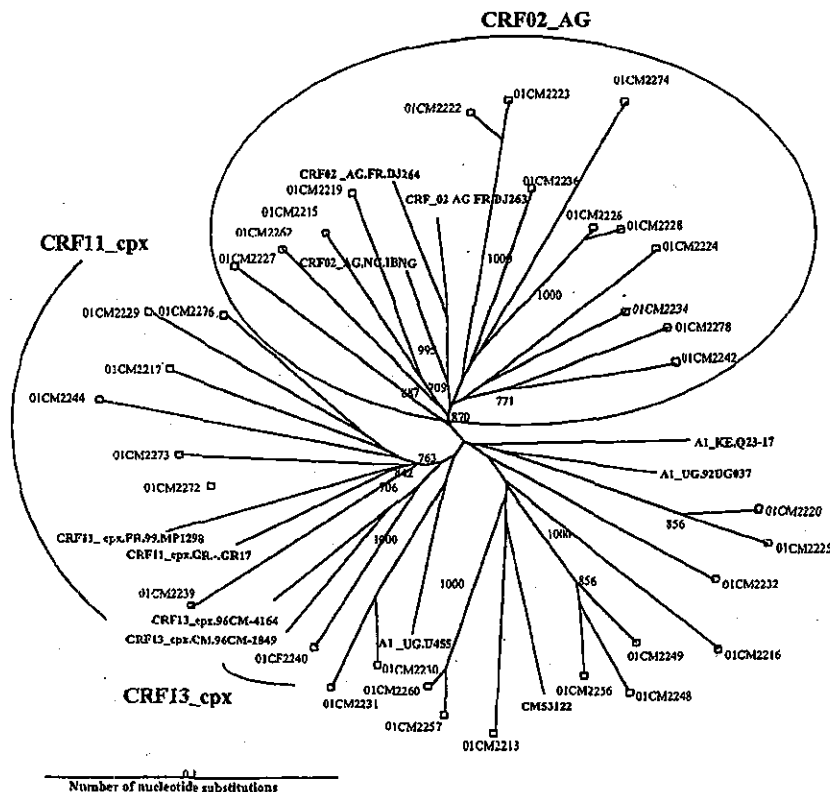
1_{HXB2}) was amplified by nested PCR with primers GP40F1 (5'-TCTTAGGAGCAGCAGGAAGCACTATGGG-3') and GP41R1 (5'-AACGACAAAGGTGAGTATCCCTGCCTAA-3') in the first round, and GP46F2 (5'-ACAATTATTGCTGGTATAGTGCAACAGCA-3') and GP47R2 (5'-TTAAACCTATCAAGCCTCCTACTATCATT-3') in the second round.²⁶ A part of *pol* gene encoding integrase (IN) (corresponding to 4493–4780 nt in HIV-1_{HXB2}) was amplified by nested PCR using primers, unipol 5 (5'-

TGGTACCAGCACACAAAGGAATAGGAGGAAA-3') and unipol 6 (5'-CCACAGCTGATCTCTGCCTTCTCTGTAATAGACC-3') in the first round, and unipol 1 (5'-AGTGGATTCATAGAAGCAGAAAGT-3') and unipol 2 (5'-CCCCTATTCCTTCCCCTTCTTTTAAAA-3') in the second round.⁶

Nested PCR was performed with an AmpliTaq Gold PCR kit (Perkin-Elmer, Foster City, CA) according to the manufacturer's instructions. Amplification was done with one

E

FIGURE 1. (continued) Phylogenetic trees based on a part of *env*-C2V3 gene (approximately 550 bp) of 40 HIV-1 strains (A), *gag* gene (approximately 460 bp) of 56 HIV-1 strains (B), *pol*-IN gene (288 bp) of 20 HIV-1 strains (C), and *env*-gp41 gene (approximately 405 bp) of 19 HIV-1 strains (D) from eastern Cameroon with reference sequences of representative subtypes/CRF. The reference sequences used in *gag*, *pol*, and *env* trees are shown in bold. The bootstrap value at each node represents the number among 1000 bootstrap replicates that support the branching order. Bootstrap value of $\geq 70\%$ or higher is shown. Brackets on the right represent the major group M subtypes. Newly derived sequences are marked with a filled square (■). Further analysis of the 34 probable subtype A sample sequences without the rest of the sample sequences is shown in part E.



cycle of 95°C for 10 minutes and 35 cycles of 95°C for 30 seconds, 45–55°C for 30 seconds, and 72°C for 1 minute with a final extension of 72°C for 10 minutes. Samples that could not be amplified with the *env*-C2V3 or the *gag*-p24 primer sets were analyzed with *env*-gp41 and *pol*-IN-specific primers. PCR amplification was confirmed by visualization with ethidium bromide staining of the gel. The amplified products were cloned with TOPO TA cloning kit (Invitrogen, Carlsbad, CA) and sequenced as previously described.⁶ At least 5 clones per sample were analyzed to obtain a consensus sequence.

Genetic Distance and Phylogenetic Analysis

The sample sequences were aligned with subtype reference sequences from the Los Alamos database by CLUSTAL W (version 1.81) with subsequent inspection and manual modification. The frequency of nucleotide substitution in each base of the sequences was estimated by the Kimura 2-parameter method.²⁷ A phylogenetic tree was constructed by the neighbor-joining method,²⁸ and its reliability was estimated by 1000 bootstrap replications. The profile of the tree was visualized with Tree View PPC version 1.6.5.

Nucleotide Sequence Accession Numbers

The nucleotide sequences in this study were submitted to Genbank and are available under the following accession numbers:

- gag*-p24 sequences: AY539583–AY539638 (n = 56)
- pol*-IN sequences: AY539639–AY539658 (n = 20)
- env*-C2V3 sequences: AY543545–AY543584 (n = 40)
- env*-gp41 sequences: AY541010–AY541028 (n = 19)

RESULTS

The epidemiologic and clinical data on the 57 HIV-1-infected individuals and genetic subtypes of HIV-1 strains identified in *gag*-p24, *env*-C2V3, *env*-gp41, and *pol*-IN are shown in Table 1.

Phylogenetic Analysis in the *env*-C2V3 Sequences

Out of the 57 samples, 40 (70.2%) were found to be positive for HIV-1 PCR with the set of *env*-C2V3 primers. The remaining 17 samples (29.8%) were negative for PCR with the *env*-C2V3 primers used in the present study. The phylogenetic tree based on *env*-C2V3 sequences (Fig. 1A) showed that 34 (85.0%) were probable subtype A, 1 (2.5%) F2, 4 (10.0%) CRF01_AE, and 1 (2.5%) U (unclassified). Thus, the majority of the samples fell into subtype A, though the clustering with subtype A reference sequences was not completely significant. We further analyzed these 34 probable subtype A sequences without the rest of the sample sequences (Fig. 1E), showing that 14 clustered with CRF02_AG reference strains, 7 with

AU3

TABLE 2. Subtypes in *gag*-p24 and *env*-C2V3 for 57 HIV-1 Strains

	<i>gag</i> -p24	<i>env</i> -C2V3	n	Total
Concordant genotypes	A	A	3	28
	F2	F2	1	
	CRF01 (A)	A	3	
	CRF01 (A)	CRF02 (A)	1	
	CRF01 (A)	CRF11 (A)	1	
	CRF02 (A)	A	2	
	CRF02 (A)	CRF02 (A)	9	
	CRF02 (A)	CRF13 (A)	1	
	CRF11 (A)	CRF02 (A)	2	
	CRF11 (A)	CRF11 (A)	5	
Discordant genotypes	B	A	2	12
	B	CRF01 (E)	1	
	B	CRF02 (A)	2	
	CRF01 (A)	CRF01.A	1	
	CRF11 (A)	CRF01 (E)	2	
	CRF13 (G)	A	1	
	CRF13 (G)	CRF01 (E)	1	
	CRF13 (G)	CRF11 (A)	1	
Incomplete or negative PCR profile between <i>gag</i> and <i>env</i>	G	U	1	17
	a*	ND	17	

a*: 16 have (+/-) and 1 has (-/-) in the *gag/env* region. ND, not detected.

CRF11_cpx, and 1 with CRF13_cpx, significantly (with 87.0, 76.3, and 100% bootstrap values, respectively).

Phylogenetic Analysis in the *gag*-p24 Sequences

All of the 57 samples but one (02CM319) were found to be positive for HIV-1 PCR with the set of *gag* primers. The phylogenetic tree based on *gag* sequences (Fig. 1B) showed that 45 (80.4%) were subtype A, 6 (10.7%) B, 1 (1.8%) F2, 3 (5.4%) CRF13_cpx, and 1 (1.8%) U. Out of the 45 subtype A samples, 17 significantly clustered with CRF02_AG reference strains (with 74.6% bootstrap value).

Phylogenetic Analysis in the *env*-C2V3 and *gag*-p24 Genes

The summary of the phylogenetic analyses based on *gag*-p24 and *env*-C2V3 for the 57 HIV-1 strains is shown in Table 2. Out of the 40 samples that could be analyzed both in *gag* and *env*-C2V3 regions, 12 (30.0%) had discordant subtype or CRF designation: 2 subtype B/A (*gag/env*), 1 B/CRF01_AE, 2 B/CRF02_AG, 1 CRF01_AE/CRF01_AE.A, 2 CRF11_cpx/CRF01_AE, 1 CRF13_cpx/A, 1

CRF13_cpx/CRF01_AE, 1 CRF13_cpx/CRF11_cpx, and 1 G/U. In 9 of these 12 strains, CRFs were involved in recombination events. Twenty-eight strains (70.0%) had the concordant subtypes or CRF designation between *gag* and *env*: 27 subtype A and 1 F2.

Phylogenetic Analysis in the *pol*-IN and *env*-gp41 Genes

Additional PCR analyses on 17 samples with 2 *gag/env* profiles, (+/-) and (-/-), were carried out with groups M, N, and O primers for *pol*-IN and *env*-gp41. Phylogenetic analyses based on *pol*-IN and *env*-gp41 sequences revealed the presence of 10 profiles: 2 U/U, 1 U/A, 1 U/CRF11_cpx, 2 CRF11_cpx/U, 3 CRF11_cpx/CRF11_cpx, 1 CRF02_AG/CRF09_cpx, 1 CRF02_AG/CRF11_cpx, 2 CRF02_AG/U, 3 CRF02_AG/CRF02_AG, and 1 group O (*pol*-IN/*env*-gp41) (Fig. 1 C and D). Thus, of the 17 HIV-1 strains negative for PCR with the *env*-C2V3 primers used, 10 (58.8%) had discordant subtype or CRF, and 7 (41.2%) had concordant one based on *gag/pol/env*-gp41 analyses (Tables 1, 2, and 3).

V3-Loop Amino Acid Sequence Analysis

The V3-loop has been shown to be important for antibody neutralization, phenotypic change, and co-receptor usage.¹⁵ Figure 2 shows the predicted V3 amino acid sequence alignment of our 40 strains that could be analyzed in *env*-C2V3 region. Thirty-five, 3, and 2 of them had V3-loop consisting of 35, 36, and 37 amino acids in length, respectively. The most common tetrameric peptide at the apex of the V3-loop (highlighted gray in Fig. 2) was GPGQ (n = 34; 85.0%). A neutral amino acid residue, such as glycine or serine, was always present at position 11, and an amino acid with a negatively charged side chain, such as asparagic acid and glutamic acid, was present at position 25. The net charge of the V3-loop was

TABLE 3. Subtypes in *pol* and *env*-gp41 for the 17 HIV-1 Strains With Incomplete and/or Negative PCR Profile Between *gag* and *env*-C2V3

<i>env</i> -gp41 (405 bp)	<i>pol</i> -Integrase (288 bp)				Total (%)
	CRF02	CRF11	U	O	
A			1		1 (5.9)
CRF02	3				3 (17.6)
CRF09	1				1 (5.9)
CRF11	1	3	1		5 (29.4)
U	2	2	2		6 (35.3)
O				1	1 (5.9)
Total (%)	7 (41.2)	5 (29.4)	4 (23.5)	1 (5.9)	17 (100)

U, unclassified; "O," group O.

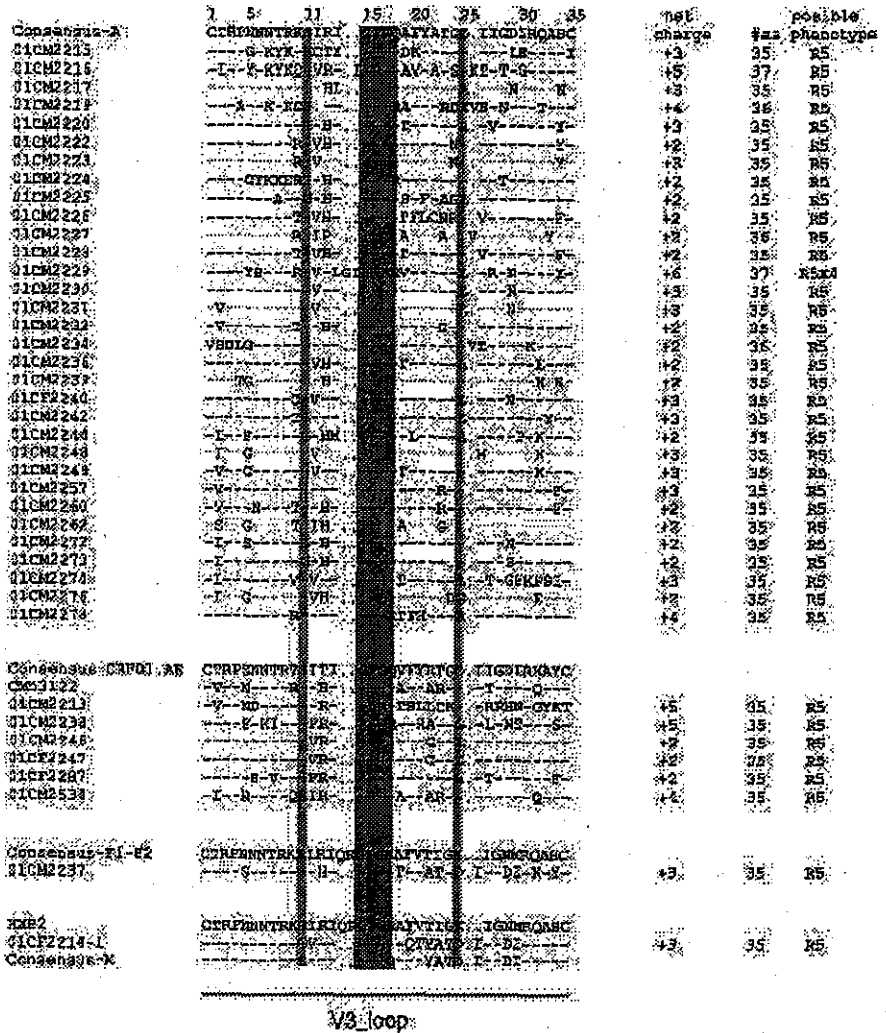


FIGURE 2. Alignment of deduced amino acid sequences of the V3-loop of HIV-1 strains from eastern Cameroon. Consensus amino acid sequences of subtype A, CRF01, subtype F2, and consensus M based on Los Alamos database are given on the top line using single letter code. Dashes and dots indicate identities and gaps, respectively. The V3-loop is indicated with a bar, and tetrameric tip of the V3-loop is highlighted in gray.

from +2 to +6 (mean net charge 2.82; 95% CI 2.49–3.15). A common feature of the HIV-1 strains with V3-loop consisting of 36 and 37 amino acids was the insertion of an amino acid residue between positions 14–15 and 25–26.

DISCUSSION

In the current study we found that a high proportion (38.6%) of HIV-1 strains in eastern Cameroon were recombinants. This is higher than those reported in urban areas (7–10%) of Cameroon.^{14,29} In our study, 4 subtype B/A (*gag/env*) recombinant strains were found in this area. This is the first report of intersubtype B/A, B/CRF01_AE, and B/CRF02_AG in Cameroon. Although CRF15_01/B (*gag/env*) and CRF03_AB have been reported in Thailand and Russia, respectively, the mosaic composition of these CRFs is different

from that of our B/A recombinants.^{30,31} However, these 4 strains did not cluster together in *env*-C2V3 tree, suggesting that these viruses have evolved from independent recombination events. In addition, our strains 01CM2213 and 01CM2534⁶ found in eastern and southern Cameroon, respectively, clustered significantly (with 80% bootstrap value) in *env*-C2V3 with a reference strain CM53122,¹³ a recombinant between CRF01_AE and subtype A found in eastern Cameroon from an unlinked epidemiologic study, suggesting a possible emergence of a second generation of CRFs. These findings strongly suggest that this part of Cameroon is a potential hotspot of HIV-1 recombination, with a likelihood of an active generation of new forms of HIV-1 variants. It would, therefore, be important to monitor the appearance of new variants in this area.

Our data in rural eastern Cameroon highlight a relatively low prevalence of CRF02_AG (29.8 and 35.0%) and relatively high prevalence of CRF11_cpx (21.4 and 17.5%) based on gag-p24 and env-C2V3 analyses, respectively. These results are in contrast with previous reports on genetic diversity in Cameroon, where the majority of env and gag subtypes A viruses are CRF02-like.^{7-15,18,29} The epidemiologic pattern of HIV-1 genetic diversity in eastern Cameroon, such as relatively low and high prevalence of CRF02_AG and CRF11_cpx, respectively, was more closely related to those of Chad and CAR than that of other geographic regions of Cameroon. A high proportion of subtype B (11%) was also found to be circulating in eastern Cameroon, and more interestingly, these subtype B sequences clustered significantly in the gag-p24 tree (with 80% of bootstrap value) with HXB2, a European subtype B reference strain. These findings could be related to the launch of a pipeline Chad-Cameroon project crossing the eastern part of Cameroon, in which not only local people but also >10,000 American and European engineers have been involved for >10 years. Further analysis of full-length genome sequences and appropriate phylogenetic analysis would be needed to clarify.

AAU As for CRF09_cpx, which sequences are available only in Los Alamos database so far, has come from independent molecular studies in Senegal and Ghana, west Africa. This is the first report of CRF09_cpx among Cameroonian isolates in rural eastern Cameroon, suggesting that it may be spreading into west-central Africa.

In the current study we used 4 sets of primers to detect highly divergent HIV-1 strains. Among the 17 samples negative for PCR with the env-C2V3 primers used, 5 (29.4%) and 6 (35.3%) were unclassified in pol-IN and env-gp41 phylogenetic analyses, respectively. Despite the great genetic diversity and the high proportion (38.6%) of recombinant strains, all of the 17 samples could be detected with the pol-IN and env-gp41 primer sets. Thus, our PCR-based approach could be a molecular screening tool for detecting highly divergent HIV-1 variants.

To predict virtual phenotype of HIV-1, 2 approaches have been used based on subtype B analyses, such as the net amino acid charge and the presence of uncharged and negatively charged amino acids at positions 11 and 25 of the V3-loop.³² In the current study, most of our strains have the V3-loop with less than +6 of net amino acid charge, suggesting that these HIV-1 strains were non-syncytium inducing viruses (NSI) using CCR5 as a coreceptor for cellular entry. In addition, of our 39 possible NSI viruses, 2 contained a positively charged amino acid residue at position 25 with a tetrameric motif consisting of GPGR at the apex of the V3-loop. The predicted phenotype did not correlate with the clinical status of our subjects, most of whom were AIDS-related complex or suspected AIDS patients. Recently, an independent study on non-B HIV-1 viruses has also revealed a lack of correlation

between predicted phenotype of isolates based on V3-based genotype and in vitro syncytium-inducing capacity of the isolates from an urban area of Cameroon.¹⁴ Further analysis will be needed for understanding of pathogenesis of non-subtype B HIV-1, and the relationship between V3-based genotype and phenotype of non-subtype B HIV-1.

In Cameroon, where multiple subtypes co-circulate, intersubtype recombination of HIV-1 strains is a common occurrence and plays a central role in active generation of new forms of HIV-1 variants. Monitoring the emergence of new strains not only in urban but also in rural areas would be very helpful for establishment of control strategies for the HIV epidemic in west-central Africa.

REFERENCES

- Robertson DL, Anderson JP, Bradac JA, et al. HIV-1 nomenclature proposal. *Science*. 2000;288:55-56.
- Gurtler LG, Hauser PH, Eberle J, et al. A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. *J Virol*. 1994;68:1581-1585.
- Simon F, Maucleere P, Roques P, et al. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med*. 1998;4:1032-1037.
- Corbet S, Muller-Trutwin MC, Vernisse P, et al. Env sequences of simian immunodeficiency viruses from chimpanzees in Cameroon are strongly related to those of human immunodeficiency virus group N from the same geographical area. *J Virol*. 2000;74:529-534.
- Zuong P, Burda S, Urbanski M, et al. HIV type 1 group M clades infecting subjects from rural villages in equatorial rain forests of Cameroon. *J Acquir Immune Defic Syndr*. 2002;31:495-505.
- Ndemi N, Yumo H, Takehisa J, et al. HIV type 1 infection in pygmy hunter gatherers is from contact with Bantu rather than from nonhuman primates. *AIDS Res Hum Retroviruses*. 2003;19:441-445.
- Heyndrickx L, Janssens W, Zekeng L, et al. Simplified strategy for detection of recombinant HIV-1 group M isolates by gag/env heteroduplex mobility assay. *J Virol*. 2000;74:363-370.
- Janssens W, Buve A, Nkengasong JN, et al. The puzzle of HIV-1 subtypes in Africa [editorial]. *AIDS*. 1997;11:705-712.
- Nkengasong JN, Janssens W, Heyndrickx L, et al. Genotypic subtypes of HIV-1 in Cameroon. *AIDS*. 1994;8:1405-1412.
- Takehisa J, Zekeng L, Ido E, et al. Various types of mixed infections in Cameroon. *Virology*. 1998;245:1-10.
- Mboudjeka I, Zekeng L, Takehisa J, et al. HIV type 1 genetic variability in the northern part Cameroon. *AIDS Res Hum Retroviruses*. 1999;15:951-956.
- Heyndrickx L, Janssens W, Ndumbe PM, et al. HIV-1 genetic diversity variability in Cameroon. *AIDS*. 2000;14:1862-1863.
- Carr JK, Torimiro JN, Wolfe ND, et al. The AG recombinant lbNG and novel strains of group M HIV-1 are common in Cameroon. *Virology*. 2001;286:168-181.
- Vergne L, Bourgeois A, Mpoungou-Ngole E, et al. Biological and genetic characteristics of HIV infections in Cameroon reveals dual group M and O infections and a correlation between SI-inducing phenotype of the predominant CRF02_AG variant and disease stage. *Virology*. 2003;310:254-266.
- Tebit DM, Zekeng L, Kaptue L, et al. Genotypic and phenotypic analysis of HIV type 1 isolates from western Cameroon. *AIDS Res Hum Retroviruses*. 2002;18:39-48.
- Janssens W, Salminen MO, Laukkanen T, et al. Near full length genome analysis of HIV type 1 CRF02_AG subtype C and CRF02_AG and subtype G recombinants. *AIDS Res Hum Retroviruses*. 2000;16:1183-1189.
- Nyambi P, Heyndrickx L, Vereecken Kalleen, et al. Predominance of infection with HIV-1 circulating recombinant form CRF02-AG in major Cameroonian cities and towns. *AIDS*. 2002;16:295-296.
- Montavon C, Vergne L, Bourgeois A, et al. Identification of a new circu-

- lating recombinant form of HIV type 1, CRF11_cpx, involving subtypes A,G,J and CRF01_AE, in Central Africa. *AIDS Res Hum Retroviruses*. 2002;18:231-236.
19. Takehisa J, Zekeng I, Eiji I, et al. Human immunodeficiency virus type 1 intergroup (M/O) recombination in Cameroon. *J Virol*. 1999;73:6810-6820.
 20. Peeters M, Liegeois F, Torimiro N, et al. Characterization of a highly replicative intergroup M/O human immunodeficiency virus type 1 recombinant isolated from Cameroonian patient. *J Virol*. 1999;73:7368-7375.
 21. National AIDS Control Program. Yaounde, Cameroon: Ministry of Public Health; 2001.
 22. Murphy E, Korber B, Georges-Courbot MC, et al. Diversity of V3 region sequences of human immunodeficiency viruses type 1 from Central African Republic. *AIDS Res Hum Retroviruses*. 1993;10:997-1006.
 23. Muller-Trutwin MC, Chaix ML, Letourneur F, et al. Increase of HIV-1 subtype A in Central African Republic. *J Acquir Immune Defic Syndr*. 1999;21:164-171.
 24. Vidal N, Koyaita D, Richard V, et al. High genetic diversity of HIV-1 strains in Chad, west Central Africa. *J Acquir Immune Defic Syndr*. 2003;33:239-246.
 25. Peeters M, Toure-Kane C and Nkengasong JN. Genetic diversity of HIV in Africa: impact on diagnosis, treatment, vaccine development and trials. *AIDS*. 2003;17:2547-2560.
 26. Yang C, Pieniazek D, Owen SM, Fridlung C, et al. Detection of phylogenetically diverse human immunodeficiency virus type 1 groups M and O from plasma by using highly sensitive and specific generic primers. *J Clin Microbiol*. 1999;37:2581-2586.
 27. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotides sequences. *J Mol Evol*. 1980;16:111-120.
 28. Saitou N and Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Bio Evol*. 1987;4:406-425.
 29. Montavon C, Toure-Kane C, Liegeois, et al. Most *env* and *gag* subtype HIV-1 viruses circulating in west and west central Africa are similar to the prototype AG recombinant virus IBNG. *J Acquir Immune Defic Syndr*. 2000;25:363-374.
 30. Tovanabutra S, Watanaveeradej V, Viputikul K, et al. A new circulating recombinant form, CRF15_01B, reinforces the linkage between IDU and heterosexual epidemics in Thailand. *AIDS Res Hum Retroviruses*. 2003;19:561-567.
 31. Liitsola K, Holm K, Bobkov A, et al. An AB recombinant and its parental HIV type 1 strains in the area of the former Soviet Union: low requirements for sequence identity in recombination. *AIDS Res Hum Retroviruses*. 2000;16:1047-1053.
 32. Jensen MA, Van't Wout AB. Predicting HIV-1 coreceptor usage with sequence analysis. *AIDS Rev*. 2003;5:104-112.

Sequence Note

Active Generation and Selection for HIV Intersubtype A/D Recombinant Forms in a Coinfected Patient in Kenya

ELIJAH M. SONGOK,^{1,2} RAPHAEL M. LWEMBE,² RUKIA KIBAYA,² KANA KOBAYASHI,¹
NICAISE NDEMBI,¹ KAYOKO KITA,¹ JOHN VULULE,² ISAO OISHI,³ FREDRICK OKOTH,²
SEIJI KAGEYAMA,¹ and HIROSHI ICHIMURA^{1,3}

ABSTRACT

To investigate the *in vivo* evolution of recombinant HIV, we followed up on a mother who was initially coinfecting with subtypes A and D in Kenya. Blood samples were obtained in 1996 and 2002, and HIV *pol* and *env* genes were amplified by PCR, cloned, sequenced, and phylogenetically analyzed. As for the 1996 sample most of the clones generated from the *pol* and *env* genes clustered either with subtypes A and D reference strains. However, two clones from the *pol* gene were found to be independent recombinants between subtypes A and D by RIP analysis, suggesting active generation of recombinant forms. As for the 2002 sample, all the clones from the *pol* gene clustered only with the subtype A reference strain, while all the *env* clones clustered only with subtype D, denoting a dominance of an A/D recombinant form. These results indicate that in patients dually infected with subtypes A and D there is an ongoing generation and selection for A/D recombinant forms.

THE GENOME OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1) is in a rapid evolutionary state. Presently, HIV has been classified into three main groups based on their genomic sequences: M (major), and two divergent groups, O (outlier) and N (non-M, non-O).¹ The vast majority of variants found worldwide and responsible for the AIDS pandemic belong to group M. Phylogenetic analysis of group M has further led to its subdivision to pure subtypes A to K and subsubtypes A1, A2, F1, and F2.² Recently, it was realized that a significant fraction of isolates is of an intersubtype form composed of genomes from two or more different subtypes. These mosaics are thought to be generated by recombination arising from coinfection of the same target cell with different viruses, leading to production of heterodimeric virions.³ They are highly likely to have arisen in geographic regions where distinct HIV subtypes co-circulate. Some of these mosaic genomes are unique and have been restricted to small clusters. However, others have spread globally infecting unlinked individuals and are now designated

as circulating recombinant forms (CRF).^{1,2} The contribution of recombinant viruses to the global AIDS pandemic is significant. The spread of CRF01-AE (CM240) in Thailand,⁴ the dominance of CRF02-AG (CRF02) in West and Central Africa,⁵ and the recent explosive outbreaks of HIV in Russia and China attributed to CRF03-AB and B/C, respectively,^{6,7} suggest that HIV recombinant variants will have profound implications in anti-HIV prevention and control strategies.

In Kenya, though pure subtypes A1 and D are the most common HIV subtypes,⁸ there has been a rise in reported cases of recombinants especially in the Western region.⁸ Samples collected between 1997 and 1998 showed that 26.8% were recombinants,⁹ while those collected 2 years later showed a more than 40% recombination.¹⁰ Expectedly, the A/D recombinant form has been the most prevalent, accounting for more than 46% of all reported recombinant cases. The increased reporting of the A/D recombinants in the country has led to speculation that in areas where A and D subtypes coexist, the inter-

¹Department of Viral Infection and International Health, Graduate School of Medical Science, Kanazawa University, Kanazawa, Japan.

²Kenya Medical Research Institute, Nairobi, Kenya.

³Japan International Cooperation Agency, The Research and Control of Infectious Diseases Project in Kenya, Nairobi, Kenya.

subtype A/D recombinant form may be preferentially selected. However, as HIV molecular studies in Kenya have mostly been cross-sectional, there has been little information on longitudinal follow-up, which has precluded analysis of their *in vivo* evolution. Demonstrating the types of HIV favored by selection mechanisms will increase our understanding of its evolution for the design of effective anti-HIV therapeutic and vaccine strategies. In this study, we report our findings on a mother in western Kenya who initially was coinfecting with HIV subtypes A and D but developed the A/D recombinant form during follow-up.

The patient involved (KS004) was in 1996 a 22-year-old antenatal clinic attendee participating in a study on the prevention of HIV mother-to-child transmission using short-course zidovudine in rural western Kenya.¹¹ She was among a group of mothers who, due to reasons earlier outlined,¹² could not comply with zidovudine use. The initial blood sample was obtained at recruitment in August 1996 and another sample was obtained 6 years later (January 2002), on a follow-up program to determine eligibility of use of full-course antiretroviral therapy. On each occasion the sample was separated into peripheral mononuclear cells using Ficoll-Paque, and genomic DNA was extracted with DNazol (GIBCO-BRL, Grand Island, NY). Polymerase chain reaction (PCR) amplification was performed on the viral *pol* and *env* regions. A part of the *pol* region encoding the reverse transcriptase gene (corresponding to 2480–3180 nt of HIV-1 HXB2) was amplified by nested PCR with

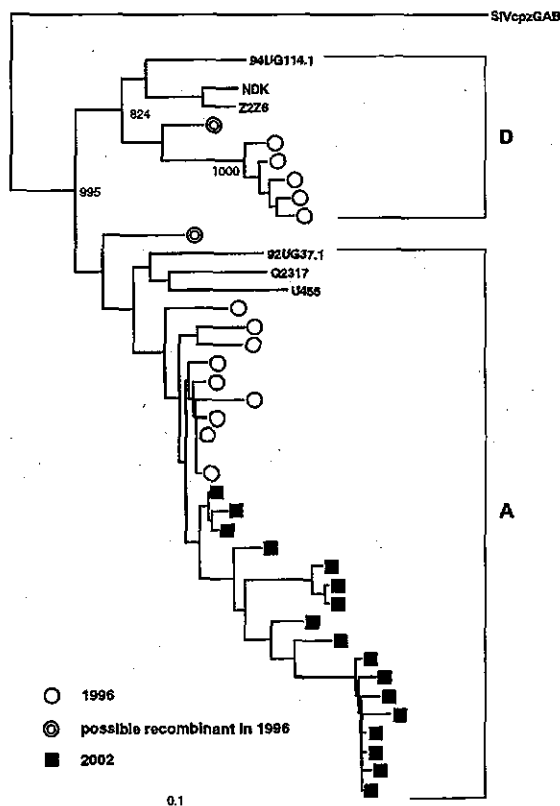


FIG. 1. Phylogenetic analysis of the *pol*-RT gene (697 bp) of HIV-1 clones isolated in 1996 (○, ⊙) and 2002 (■) from an HIV-infected mother (KS004) in Kenya.

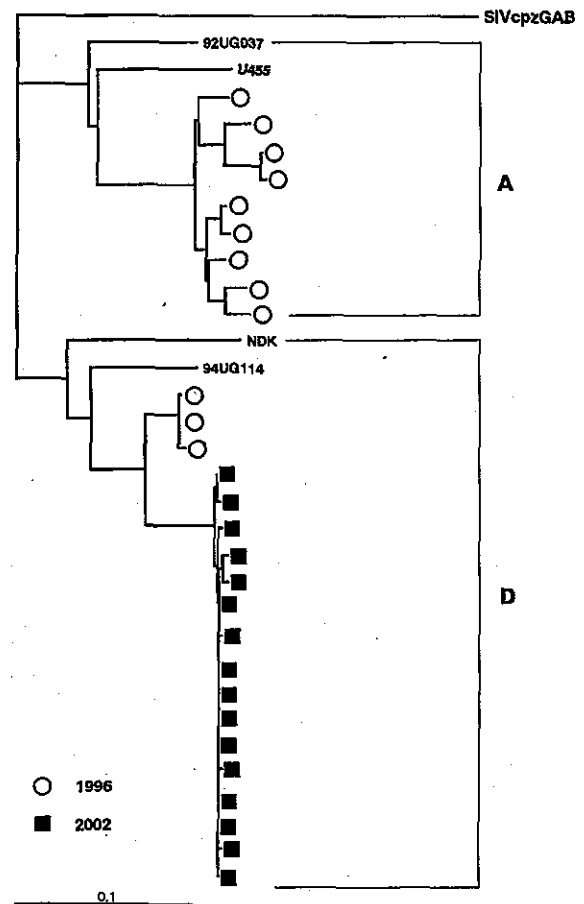


FIG. 2. Phylogenetic analysis of the *env*-C2V3 gene (550 bp) of HIV-1 clones isolated from KS004 in 1996 (○) and 2002 (■).

primer pairs RT18 (5'-GGAAACCAAAAATGATAGGGGG-AATTGGAGG-3') and KS104 (5'-TGACTTGCCCAATT-TAGTTTCCCACTAA-3') in the first round and KS101 (5'-GTAGGACCTACACCTGTTCAACATAATTGGAAG-3') and KS102 (5'-CCCATCCAAAGAAATGGAGGAGGTTCTTTCTGATG-3') in the second round. Cycling conditions included a hot start at 95°C for 10 min, then 35 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, and extension at 72°C for 1 min and a final extension at 72°C for 10 min. For the amplification of part of the *env* gene (corresponding to 6954–7504 nt of HIV-1 HXB2) primer pairs M5 (5'-CCAATTCCCATA CATTATTGTGCC CAGCTGG-3') and M10 (5'-CCAATTGTCCC TCATATCTCCTCCTCC-AGG-3') were used for the first round and primer pairs M3 (5'-GTCAGCACAGTACAATGIACACATGG-3') and M8 (5'-TCCTTGGATGGGAGGG GCATACATTGC-3') for the second round. Cycling conditions for both fragments were similar and included a hot start at 95°C for 10 min, then 35 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, and extension at 72°C for 1 min, and a final extension of 72°C for 10 min. The PCR products arising from four independent PCR reactions (approximately 697 bp for *pol* and 550 bp for *env*) were separately cloned by using a pCRII vector (In-

sults suggest that in 1996 the patient was coinfecting with a likelihood of various viral variants, such as either pure subtypes A and D, a combination of the pure subtypes and recombinants, or various types of recombinants. However, by 2002 only one variant, an A/D recombinant, had been selected for dominance.

The predicted amino acid changes in the V3 loop are shown in Figure 4. The V3 loops of the subtype D clones isolated in 1996 and those isolated in 2002 had 34 amino acids, which is consistent with other subtype D clones from the region.¹ However, there was a predicted shift in coreceptor usage from CCR5 to CXCR4 based on the change in the net amino acid charge, from +2 in 1996 to +6 in 2002.¹³ Similar amino acid changes were observed at position 7 and a whole peptide change, RQGTHI to TIYSYR, from positions 9 to 14.

Overall, isolates obtained from the 1996 sample showed greater diversity in the subtype A (both *pol* and *env*) than in the subtype D sequences. But those obtained in 2002 were almost a clonal line of *env* sequences with very little diversity. This homogeneity in *env* sequences from 2002 may be due to a waning of the host immune pressure, though the possibility of an artifact arising from our PCR and cloning procedures cannot be excluded. We are currently analyzing the viral RNA to compare the circulating virus with the proviral DNA to confirm these observations.

The emergence and dominance of the A/D recombinant form confirm not only that recombination *in vivo* occurs in dually infected patients, but also that particular types of recombinants may be preferred by selection processes over the parental subtypes. A recent report from the same location in Kenya has shown a preference for vertical transmission of subtype D containing *gag* and *env* fragments over subtype A.¹⁴ Similarly, in a large study cohort in neighboring Uganda, there have been reports of faster progression to AIDS among cases infected with HIV-1 subtype D than those with subtype A in the *env* region.¹⁵ The preferred selections for *env* D over *env* A in our study case may support the hypothesis that viruses containing subtype D sequences in the *env* region have better fitness capacity *in vivo*.

Intersubtype recombination may be a common occurrence in Kenya and plays a central role in the emergence of new HIV variants. However, for the design of effective anti-HIV strategies, there is a need for more studies in areas where multiple HIV variants cocirculate to establish why some subtypes or recombinants have an ability to subsequently "win" over the rest for dominance.

GeneBank accession numbers: deposit of samples to GeneBank in process.

ACKNOWLEDGMENTS

The authors wish to thank all counterparts of the KEMRI-JICA Project for their services. Special thanks to Dr. Takahashi Kurimura and Dr. Davy Koech for their technical advice. This work was supported in part by International Scientific Research Program (Grant 14256005) from Monbu-Kagakusho (Japan Ministry of Education and Science) and by The Research and Control of Infectious Diseases Project in Kenya, JICA. This work is published with the permission of the Director, Kenya Medical Research Institute.

REFERENCES

1. Kuiken C, Foley B, Hahn B, *et al.*: *HIV Sequence Compendium 2001*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, 2001.
2. Pecters M and Sharp PM: Genetic diversity of HIV-1: The moving target. *AIDS* 2000;14(Suppl. 3):S129-139.
3. Robertson DL, Sharp PM, McCutchan FE, *et al.*: Recombination in HIV-1. *Nature* 1995;375:124-126.
4. Gao F, Robertson DL, Morrison SG, *et al.*: The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype A/E recombinant of African origin. *J Virol* 1996;70:7013-7029.
5. Papathanasopoulos MA, Hunt GM, and Tiemessen CT: Evolution and diversity of HIV-1 in Africa—a review. *Virus Genes* 2003; 26:151-163.
6. Littola K, Tashkinova T, Laukkanen G, *et al.*: HIV-1 genetic subtype A/B recombinant strain causing explosive epidemic in injecting drug users in Kalingrad. *AIDS* 1998;12:1907-1919.
7. Piyasirisilp S, McCutchan FE, Carr JK, *et al.*: A recent outbreak of human immunodeficiency virus type 1 infection in Southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC recombinant. *J Virol* 2000;74:11286-11295.
8. Songok EM, Lihana RW, Kiptoo MK, *et al.*: Identification of *env* CRF-10 among HIV variants circulating in rural Western Kenya. *AIDS Res Hum Retroviruses* 2003;19:161-165.
9. Heyndrickx L, Janssens W, Zekeng L, *et al.*: Simplified strategy for detection of recombinant human immunodeficiency virus type 1 group M isolates by *gag/env* heteroduplex mobility assay. *J Virol* 2000;74:363-370.
10. Dowling WE, Kim B, Marson CJ, *et al.*: Forty-one near full-length HIV sequences from Kenya reveal an epidemic of subtype A and A-containing recombinants. *AIDS* 2002;16:1809-1820.
11. Songok EM, Kakimoto K, Genga I, *et al.*: Prenatal short course zidovudine reduces mortality in children born to HIV-1 positive mothers in rural Kenya. *J Infect Dis* 2001;183:1540-1541.
12. Songok EM, Fujiyama Y, Tukei PM, *et al.*: The use of short course zidovudine to prevent perinatal transmission of human immunodeficiency virus in rural Kenya. *Am J Trop Med Hyg* 2003;69: 8-13.
13. McDonald RA, Chang G, Michael NL: Relationship between V3 genotype, biologic phenotype, tropism and co-receptor use for primary isolates of human immunodeficiency virus type 1. *J Hum Virol* 2001;4:179-187.
14. Chunfu Y, Ming L, Newman DR, *et al.*: Genetic diversity of HIV-1 in western Kenya: Subtype specific differences in mother-to-child transmission. *AIDS* 2003;17:1667-1674.
15. Kaleebu P, French N, Mahe C, *et al.*: Effect of human immunodeficiency virus (HIV) type 1 envelope subtypes A and D on disease progression in a large cohort of HIV-1 positive persons in Uganda. *J Infect Dis* 2002;185:1244-1250.

Address reprint requests to:

Hiroshi Ichimura
Department of Viral Infection and International Health
Graduate School of Medical Science
Kanazawa University
13-1 Takara-machi
Kanazawa 920-8640 Japan

E-mail: ichimura@med.kanazawa-u.ac.jp

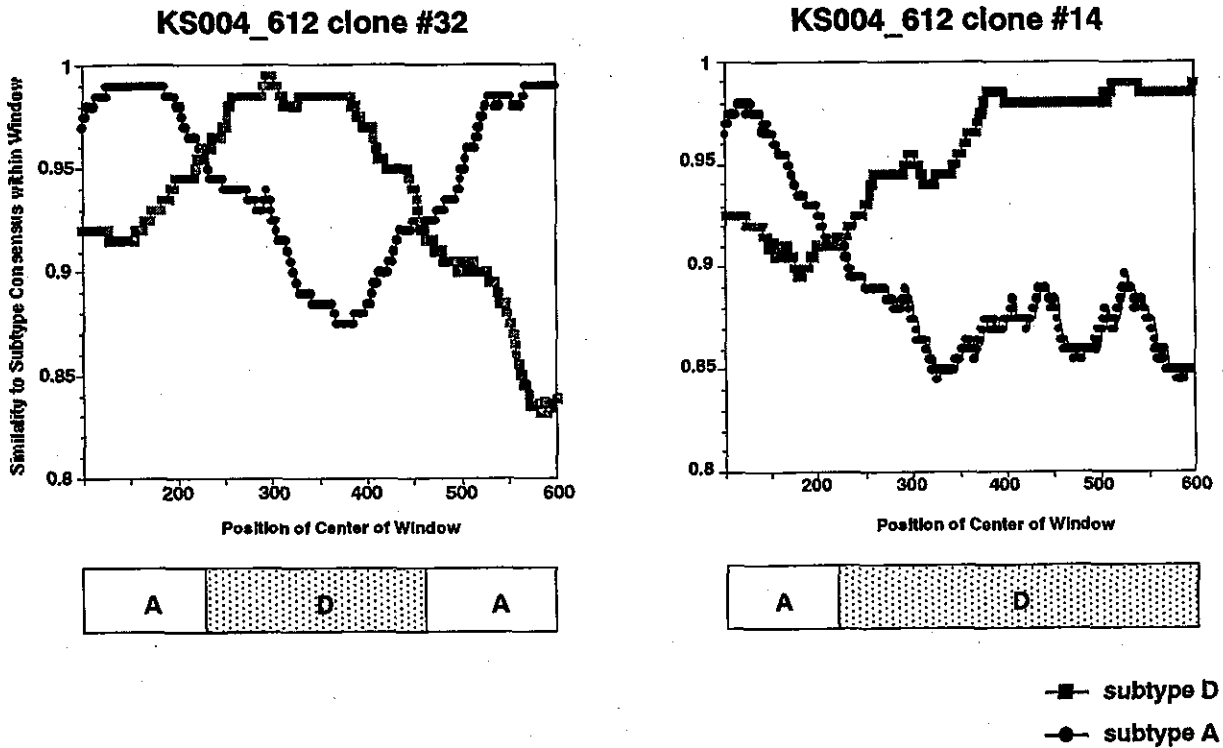


FIG. 3. RIP output and schematic representation of the mosaic structure of the *pol*-RT gene (697 bp) of HIV-1 clones (KS004_612 clones #14 and #32) isolated from KS004 in 1996.

vitrogen, Carlsbad, CA) and plasmid DNA for sequencing prepared by Quantum miniprep kit (Bio-Rad Life Sciences, Hercules, CA). DNA sequencing was carried out by dye-deoxy terminator chemistry on an ABI 310 automatic sequencer (Applied Biosystems, Foster City, CA). At least 12 plasmid clones were sequenced from each fragment and from samples collected in both of the years to obtain the consensus sequence. DNA sequences were aligned by using CLUSTAL W, version 1.81, with subsequent inspection and manual modification. Phylogenetic trees were constructed by the neighbor-joining method and its reliability was estimated by 1000 bootstrap replications. All alignments were gap stripped for the generation of trees. Resulting trees were visualized using Treeview, version 1.65. To

determine recombination breakpoints, the recombinant identification (RIP) analysis was performed.¹

The outcome of the phylogenetic analysis of the reverse transcriptase *pol* gene region (697 bp) of representative clones of both the 1996 and 2002 samples is shown in Figure 1. Clones derived from samples collected in 1996 clustered with representative sequences from either subtype A or D in both the *pol* (Fig. 1) and *env* gene (Fig. 2). Two clones generated from the *pol* gene of the 1996 sample did not cluster with any reference strains, and RIP analysis revealed that they were independent recombinant forms between subtypes A and D (Fig. 3). Clones isolated from the 2002 sample all clustered within the subtype A sequences in the *pol* and subtype D in *env* gene. These re-

AA POSITION	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	Net charge	
Consensus D	C	T	R	P	Y	N	N	T	R	Q	S	T	H	I	G	P	G	Q	A	L	Y	T	T	K	.	I	I	G	D	I	R	Q	A	H	C	
KS004_612-04	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	R	-	W	-	-	-	E	.	V	-	-	-	-	-	-	-	-	+2	
KS004_612-05	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	R	-	W	-	-	-	E	.	V	-	-	-	-	-	-	-	-	+2	
KS004_612-10	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	R	-	W	-	-	-	E	.	V	-	-	-	-	-	-	-	-	+2	
KS004_201-01	-	-	-	-	-	-	K	-	T	I	Y	S	Y	R	-	-	-	R	-	W	-	-	-	T	.	R	-	-	-	-	-	-	-	-	+6	
KS004_201-02	-	-	-	-	-	-	K	-	T	I	Y	S	Y	R	-	-	-	R	-	W	-	-	-	T	.	R	-	-	-	-	-	-	-	-	+6	
KS004_201-03	-	-	-	-	-	-	K	-	T	I	Y	S	Y	R	-	-	-	R	-	W	-	-	-	T	.	R	-	-	-	-	-	-	-	-	+6	
KS004_201-05	-	-	-	-	-	-	K	-	T	I	Y	S	Y	R	-	-	-	R	-	W	-	-	-	T	.	R	-	-	-	-	-	-	-	-	+6	

FIG. 4. Predicted amino acid changes in the V3 loop of KS004 clones isolated in December 1996 (KS004_612) and in January 2002 (KS004_201).