JOURNAL OF VIROLOGY, Jan. 2000, p. 363–370 Vol. 74, No. 1
0022-538X/00/$04.00+0

Simplified Strategy for Detection of Recombinant Human Immunodeficiency Virus Type 1 Group M Isolates by *gag*/*env* Heteroduplex Mobility Assay

LEO HEYNDRICKX,[1] WOUTER JANSSENS,[1,2]* LÉOPOLD ZEKENG,[3] ROSEMARY MUSONDA,[4] SÉVERIN ANAGONOU,[5] GERT VAN DER AUWERA,[1] SANDRA COPPENS,[1] KATLEEN VEREECKEN,[1] KO DE WITTE,[1] RIAN VAN RAMPELBERGH,[1] MAINA KAHINDO,[6] LINDA MORISON,[7] FRANCINE E. MCCUTCHAN,[8] JEAN K. CARR,[8] JAN ALBERT,[9] MAX ESSEX,[10] JAAP GOUDSMIT,[11] BIRGITTA ASJÖ,[12] MIKA SALMINEN,[13] ANNE BUVÉ,[1] STUDY GROUP ON HETEROGENEITY OF HIV EPIDEMICS IN AFRICAN CITIES,† AND GUIDO VAN DER GROEN[1]

Department of Microbiology, Institute of Tropical Medicine, Antwerp,[1] and The Flanders Interuniversity Institute for Biotechnology (VIB), Zwijnaarde,[2] Belgium; Laboratoire de Santé Hygiène Mobile, Ministère de la Santé, Yaoundé, Cameroon[3]; Tropical Diseases Research Centre, Ndola, Zambia[4]; Programme National de Lutte contre le SIDA, Cotonou, Bénin[5]; National AIDS/STD Control Programme, Nairobi, Kenya[6]; London School of Hygiene & Tropical Medicine, Department of Infections and Tropical Diseases, London, United Kingdom[7]; Henry M. Jackson Foundation, Rockville, Maryland[8]; Swedish Institute for Infectious Disease Control, Karolinska Institute, Stockholm, Sweden[9]; Harvard Institute and the Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts[10]; Department of Human Retrovirology, Amsterdam, The Netherlands[11]; National Centre for Research in Virology, University of Bergen, Bergen, Norway[12]; and Department of Chronic Viral Infections, National Public Health Institute, Helsinki, Finland[13]

Received 10 May 1999/Accepted 22 September 1999

We developed a heteroduplex mobility assay in the *gag* gene (*gag* HMA) for the identification of group M subtypes A to H. The assay covers the region coding for amino acid 132 of p24 to amino acid 20 of p7 (according to human immunodeficiency virus type 1 [HIV-1] ELI, 460 bp). The *gag* HMA was compared with sequencing and phylogenetic analysis of an evaluation panel of 79 HIV-1 group M isolates isolated from infected individuals from different geographic regions. Application of *gag* HMA in combination with *env* HMA on 252 HIV-1-positive plasma samples from Bénin, Cameroon, Kenya, and Zambia revealed a high prevalence of a variety of intersubtype recombinants in Yaoundé, Cameroon (53.8%); Kisumu, Kenya (26.8%); and Cotonou, Bénin (41%); no recombinants were identified among the samples from Ndola, Zambia. The AG_{IbNG} circulating recombinant form, as determined by *gag* HMA, was found to be the most common intersubtype recombinant in Yaoundé (39.4%) and Cotonou (38.5%). Using a one-tube reverse transcriptase PCR protocol, this *gag* HMA in combination with *env* HMA is a useful tool for rapidly monitoring the prevalence of the various genetic subtypes as well as of recombinants of HIV-1. Moreover, this technology can easily be applied in laboratories in developing countries.

The Human immunodeficiency virus type 1 (HIV-1) exhibits an extremely high genetic variation, which is driven by a high error rate of the reverse transcriptase (RT), the presence of viral RNA as a dimer (recombination occurs frequently during reverse transcription) (29), the high turnover of HIV-1 in vivo (14), and selective immune responses.

By genetic analyses, HIV strains collected from around the world have been shown to have substantial diversity (19). Two types have been characterized in humans: HIV-1, the predominant HIV type throughout the world, and HIV-2, which is less widespread and still primarily found in West Africa. Three different HIV-1 groups are distinguished: group M (major), which is globally prevalent; group O (outlier) (5); and group N (non-M/non-O) (27). HIV-1 groups and subtypes are unevenly distributed throughout the world (2, 16, 17, 23). Over the years the definition of group M subtypes has been adapted in the light of emerging recombinants. Representatives of different "pure" (nonrecombinant) subtypes A, B, C, D, F, G, H, and J and circulating recombinant forms (CRFs) AE_{CM240}, AG_{IbNG}, AGI_{CY032}, and AB_{Kal153}, were proposed based on near-full-length genome analysis, as determined by the HIV Sequence

* Corresponding author. Mailing address: Department of Microbiology, Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerp, Belgium. Phone: 32-3-247-63-28. Fax: 32-3-247-63-33. E-mail: wjanssens@itg.be.

† Members of the Study Group on Heterogeneity of HIV Epidemics in African Cities include A. Buvé (coordinator), M. Laga, E. Van Dyck, W. Janssens, and L. Heyndrickx (Institute of Tropical Medicine, Antwerp, Belgium); M. Caraël (UNAIDS); S. Anagonou (Programme National de Lutte contre le SIDA, Bénin); M. Laourou (Institut National de Statistiques et d'Analyses Economiques, Bénin); L. Kanhonou (Centre de Recherche en Reproduction Humaine et en Démographie, Bénin); L. Zekeng (Laboratoire de Santé Hygiène Mobile, Cameroon); E. Akam and M. de Loenzien (Institut de Formation et de Recherche en Démographiques, Cameroon); S.-C. Abega (Université Catholique d'Afrique Centrale, Cameroon); M. Kahindo (formerly National AIDS/STD Control Programme, Kenya); J. Chege and N. Rutenberg (The Population Council, Nairobi); V. Kimani (Department of Community Health, University of Nairobi); R. Musonda, T. Sukwa, and F. Kaona (Tropical Diseases Research Centre, Zambia); B. Auvert and E. Lagarde (INSERM U88, Paris, France); N. J. Robinson (formerly INSERM U88, Paris, France); B. Ferry and N. Lydié (Centre Français sur la Population et le Développement, Paris, France); and R. Hayes, L. Morison, H. Weiss, and J. Glynn (London School of Hygiene and Tropical Medicine, London, United Kingdom).

TABLE 1. *gag* HMA reference panel[a]

Isolate	AG_{IbNG}	A	B	C	D	AE_{CM240}	F	G	H
1	LBV23-10	VI310	TB132	DJ259	K31	TN238	BZ162	LBV21-7	VI525
2	DJ258	CI4	UG280	VI313	VI203	CA10[c]	VI174	RU520	VI991
3	CI51	VI57	PIC63[b]	UG268	VI205		VI69	RU131	VI997
4		K29	PIC335[b]		VI761[c]				
5					PIC49[c]				

[a] Unless otherwise indicated, all clones have a 1,500-bp insert (20).
[b] *gag* fragment (830 bp); nucleotides 822 to 1652 according to ELI (accession number K03454).
[c] *gag* fragment (1,216 bp); nucleotides 436 to 1652 according to ELI.

Database (HSB; J. K. Carr, B. T. Foley, T. Leitner, M. Salminen, B. Korber, and F. McCutchan [http://hiv-web.lanl.gov/]). These are needed as a framework to identify new subtypes and intersubtype recombinants, which are an important source of HIV-1 genetic variation. The most prevalent CRFs are AE_{CM240} and AG_{IbNG}. AE_{CM240} is common in the Central African Republic, Thailand, and other Asian countries (11). In Nigeria AG_{IbNG} recombinants (named after the first isolate from Ibadan, Nigeria) were identified (4, 15, 22). Additional AG_{IbNG}-like full-length sequences were also obtained from

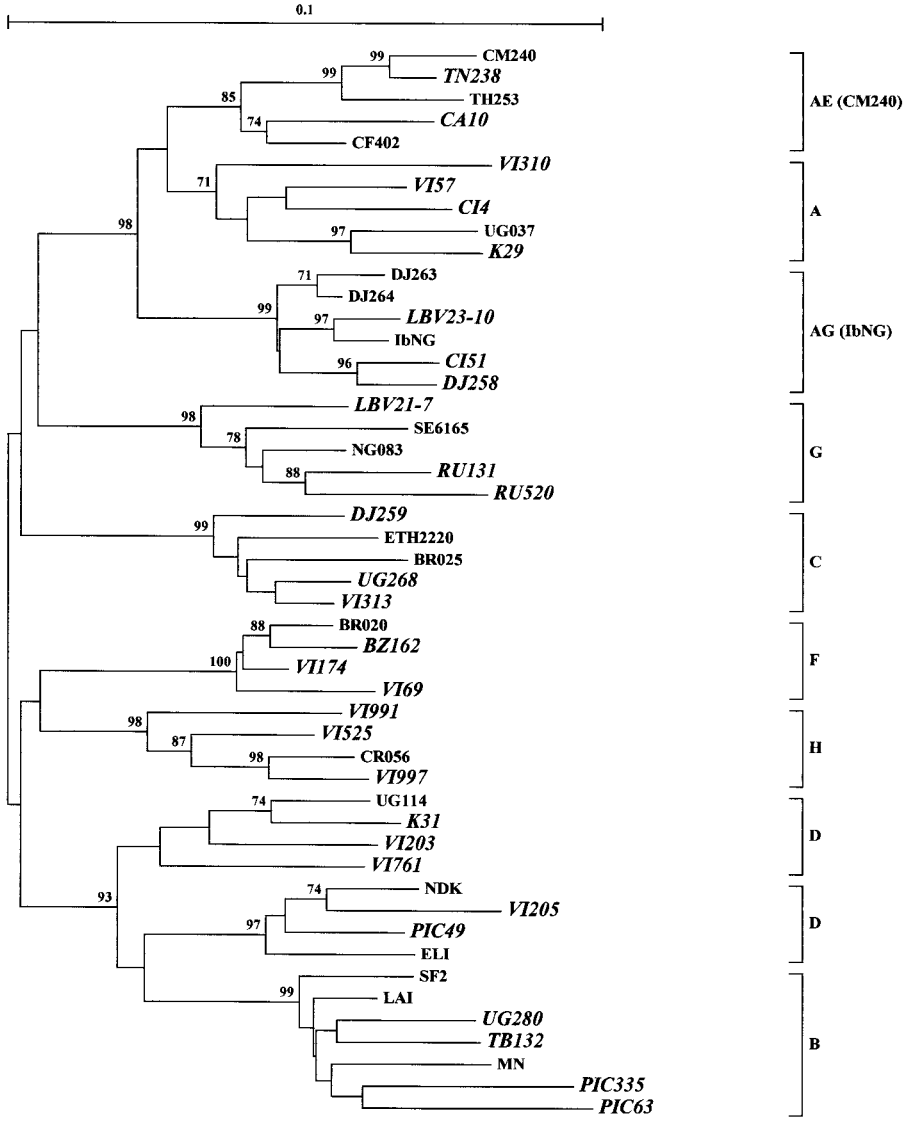

FIG. 1. Phylogenetic tree based on a *gag* gene region (460 bp; nucleotides 1123 to 1589; according to HIV-1 ELI) encoding amino acid 132 of p24 to amino acid 20 of p7. References of the *gag* HMA panel are indicated in italic. A total of 2,000 bootstrap samples were analyzed. Bootstrap values are given in percentages at the internodes if they exceed the 70% level. The distance between two sequences is obtained by summing the lengths of the connecting horizontal branches by using the scale on top. The tree is rooted arbitrarily.

TABLE 2. *gag* subtypes[a]

Origin	No. of subtypes								
	a/b[b]	A	AE_{CM240}	AG_{IbNG}	C	D	F	G	U[c]
ZM	61/61	1			59				1[d]
CM	65/66	24	1	26		3	1	9	1[e]
KE	82/86	50			9	19			4[f]
BJ	34/39	11		15				8	
Total	242/252	86	1	41	68	22	1	17	6
(%)	(96)	(35.5)	(0.4)	(16.9)	(28.1)	(9.1)	(0.4)	(7)	(2.5)

[a] Subtyping by HMA, sequencing, and phylogenetic analysis of a fragment of ca. 460 bp covering a *gag* gene region coding for amino acid 132 of p24 to amino acid 20 of p7 (according to HIV-1 ELI; accession number K03454).
[b] a, number of *gag* HMA PCR positive samples; b, number of tested samples.
[c] U, not classified by *gag* HMA.
[d] Subtype C sample.
[e] Sample remains unclassifiable by sequencing and phylogeny.
[f] Subtypes by sequencing and phylogeny: A, $n = 1$; D, $n = 2$; dual D+C, $n = 1$.

Djibouti (4) and Côte d'Ivoire (per HSB). Reanalysis of partial *env* sequences from Côte d'Ivoire (9) and Cameroon (28), initially documented as subtype A, indicates that strains of HIV-1 similar to AG_{IbNG} are the most common form of subtype A in Côte d'Ivoire and Cameroon (per HSB and J. K. Carr, personal communication).

To date, there have been few systematic large-scale attempts to characterize HIV isolates from different parts of the world. Thus, our knowledge about the distribution of HIV strains in different populations and about changes in that distribution over time is rather limited. *env* heteroduplex mobility assay (HMA) is much less cumbersome and less expensive than nucleic acid sequencing and phylogenetic analysis (6, 7). Since there is a strong correlation between the subtyping result obtained by HMA and that obtained by sequencing and phylogenetic analysis (6, 13), HMA has been introduced by UNAIDS in several developing countries as a tool for monitoring subtype distribution. Adding *gag* HMA to *env* HMA would allow one to provide a "minimum estimate" of the prevalence of the various genetic subtypes as well as some of the recombinants of HIV-1.

MATERIALS AND METHODS

Reference panel. A panel of 23 plasmids containing the *gag* gene of HIV-1 strains belonging to group M subtypes A to H was available (20). This panel was extended by adding plasmids containing part of the *gag* gene of CA10 (CRF AE_{CM240}), VI991 and VI997 (subtype H). Following evaluation of this panel, *gag* HMA reference plasmids for subtype B (PIC63 and PIC335) and subtype D (VI761 and PIC49) were added (Table 1).

Evaluation panel. There were 10 supernatants from cultures for which full-length HIV-1 sequences were available: subtypes A ($n = 6$; SE6594, SE7253, SE7535, SE8131, SE8538, and SE8891), AG_{IbNG} ($n = 1$; SE7812), G ($n = 1$; SE6165), and J ($n = 2$; SE9173 and SE9280) (per HSB). There were five supernatants from cultures for which the *gag*/*env* subtype was determined by sequencing and phylogenetic analysis: subtype F ($n = 5$; 93R26, 96R85, 96R95, 97R102, and 97R110) (E. Op de Coul, R. van den Burg, B. Asjö, A. Cupsa, R. Pascu, C. Usein, J. Goudsmit, and M. Cornelissen, unpublished data). There were 4 plasmids containing full-length HIV-1 sequences of subtype C ($n = 4$; 96BW01, 96BW04, 96BW05, and 96BW15) (26); 13 plasmids containing (part of) the *gag* gene of subtypes A ($n = 7$; K98, K7, K88, K112, K89, VI415, and VI32), D ($n = 3$; UG274, UG270, and SE365), AE_{CM240} ($n = 1$, CM240), and A/G ($n = 2$; CI20 and CI32) (20); and 11 peripheral blood mononuclear cell cultures for which the *gag*/*env* subtypes were A/A ($n = 2$; VI537 and VI1139), A/D ($n = 1$; VI1308), B/B ($n = 1$; VI1663), B/ND ($n = 2$; PH136 and PH153), C/C ($n = 2$; ZM18 and ZM20), AE_{CM240} ($n = 2$; CM235 and CM243), and F/F ($n = 1$; BZ126) (3, 20, 21). Plasma of 36 HIV-1 group M (A to H) infected patients, from different geographic regions, for which the *gag*/*env* subtype was determined by sequencing and phylogeny, was also tested: subtypes A/A ($n = 2$; PIC349 and PIC382), A/D ($n = 1$; PIC25), AE_{CM240} ($n = 4$; PIC6, PIC807, PIC231, and PIC1047), A/G ($n = 2$; PIC436 and VI1197), B/B ($n = 10$; PIC63, PIC55, PIC135, PIC1626, PIC1852, PIC271, PIC1622, PIC143, PIC335, and PIC1585), B/A ($n = 1$; PIC364), C/C ($n = 5$; VI 882, VI1044, VI1144, VI1233, and PIC146), D/D ($n = 4$; VI205, VI761, PIC797, and PIC1018), D/F ($n = 3$; PIC49, PIC179, and PIC885), F/F ($n = 1$; PIC132), G/A ($n = 1$; PIC201), and H/H ($n = 2$; VI997 and PIC450).

Study subjects. Between July 1997 and January 1998 serum samples and corresponding data were collected from roughly 1,000 men and 1,000 women, who were randomly selected from the general population, as well as 300 commercial sex workers, in each of four sub-Saharan African cities (Cotonou, Bénin [BJ]; Yaoundé, Cameroon [CM]; Kisumu, Kenya [KE]; and Ndola, Zambia [ZM]). Among the HIV-positive subjects, plasma samples were analyzed from 39 HIV-1-infected individuals from the general population; among HIV-infected commercial sex workers from Cotonou, 66 were from Yaoundé, 61 were from Ndola, and 86 were from Kisumu.

TABLE 3. *env* subtypes[a]

Origin	No. of subtypes									
	a/b[b]	A[c]	C	D	AE_{CM240}	F	G	H	U[d]	MI[e]
ZM	61/61		59						1	1
CM	65/66	48		4	2	4	5	1	1	
KE	86/86	48	7	26			1		2	2
BJ	39/39	29					10			
Total	251/252	125	66	30	2	4	16	1	4	3
(%)	(99.6)	(49.8)	(26.3)	(12.0)	(0.8)	(1.6)	(6.4)	(0.4)	(1.6)	(1.2)

[a] Subtyping by HMA (ES7-ES8 fragment) (6) or by sequencing and phylogenetic analysis of a C2V3 encoding *env* fragment if samples could not be amplified for the ES7-ES8 fragment or could not be classified by HMA.
[b] a, number of *env* PCR positive samples; b, number of tested samples.
[c] A, subtype A and AG_{IbNG} could not be differentiated by *env* HMA.
[d] U, unclassifiable by HMA and C_2V_3 sequence analysis.
[e] MI, multiple infections. Three cases of dual infection were found: ZM, A+C ($n = 1$); KE, A+D ($n = 2$).

RNA extractions and RT-PCR. RNA extractions were performed as previously described (1). One-tube RT-PCR (Access RT-PCR; Promega, Leiden, The Netherlands) was performed according to the manufacturer's recommendations.

(i) *gag* one-tube RT-PCR. The primers were H1G777 (5′-TCACCTAGAAC TTTGAATGCATGGG-3′) and H1P202 (5′-CTAATACTGTATCATCTGCTC CTGT-3′), and the cycle protocol was 45 min at 48°C (cDNA reaction), followed by 2 min at 94°C and 40 cycles for 30, 30, and 90 s at, respectively, 94, 50, and 68°C; and 1 cycle for 7 min at 68°C. For nested PCR, the primers were H1Gag1584 (5′-AAAGATGGATAATCCTGGG-3′) and g17 (5′-TCC ACATTTCCAACAGCCCTTTTT-3′), and the cycling conditions were 1 cycle for 2 min at 94°C; 35 cycles for 30, 30, and 60 s at, respectively, 94, 50, and 72°C; and 1 cycle for 7 min at 72°C. A larger fragment (830 bp) containing the entire *gag* HMA fragment was amplified by using H1G822 (5′-GCTTTCAG CCCAGAAGTAATACC-3′) and H1GHMA1317 (5′-CCAAATTCTCCCTA AAAAATTAGCCT-3′) during the nested PCR and by using the same cycling conditions as for H1Gag1584 and g17.

(ii) *env* one-tube RT-PCR. The primers ED5 and ED12 (6) were used, and the cycle protocol was 45 min at 48°C (cDNA reaction), 1 cycle for 2 min at 94°C; 40 cycles for 30, 30, and 120 s at, respectively, 94, 55, and 68°C; and 1 cycle for 7 min at 68°C. For nested PCR, the primers were ES7 and ES8 (6).

(iii) *pol* one-tube RT-PCR. The primers H1P4235 (10) and H1P5155as (5′-C TCTGTGGCCCCTGGTCTTCT-3′) were used. The cycle protocol was 45 min at 48°C (cDNA reaction), followed by 2 min at 94°C and 40 cycles for 30, 30, and 90 s at, respectively, 94, 55, and 68°C; and 1 cycle of 7 min at 68°C. For nested PCR the primers H1P4327 (10) and H1P5128as (5′-CTCTTTCCATCTGTCTT CTGCTA-3′) were used, and the cycling conditions were 1 cycle for 2 min at 94°C; 35 cycles for 30, 30, and 60 s at, respectively, 94, 50, and 72°C; and 1 cycle for 7 min at 72°C.

Sequence analysis. Sequencing of both DNA strands was performed by cycle sequencing and 5′-fluorescein isothiocyanate (FITC)-labeled primers on an automated laser fluorescence sequencer (Amersham Pharmacia Biotech, Roosendaal, The Netherlands). Sequences were aligned with CLUSTAL W (30), and the resulting alignments were refined manually with the dedicated comparative sequence editor (8). The software package TREECON was used for distance calculations (Jukes and Cantor), tree construction (neighbor-joining method), and bootstrap analysis (31). The HIV-1 *gag* nucleotide sequence data were deposited in the EMBL, GenBank, and DDBJ nucleotide sequence databases under the following accession numbers: AF184346 to AF184587.

Heteroduplex formation and mobility analysis. HMA was largely performed as described by Delwart et al. (6, 7). Briefly, the following methodology was used. Heteroduplex molecules were obtained by mixing 4.5 μl from two second-round PCRs and adding 1 μl of 10× annealing buffer (1 M NaCl, 100 mM Tris-HCl [pH 7.8], 20 mM EDTA). The DNA fragments were denaturated at 94°C for 2 min and reannealed by rapid cooling on wet ice.

Electrophoresis was done on a 5% polyacrylamide gel (29:1 acrylamide:bisacrylamide) that included 20% urea in 1× TBE buffer (88 mM Tris-borate, 89 mM boric acid, 2 mM EDTA) at 250 V for 2.5 h. Detection of heteroduplexes was done by staining with ethidium bromide and visualization under UV light.

In order to better discriminate between subtype A, CRF AE_{CM240}, and CRF AG_{IbNG}, the gel composition was altered by adding 30% urea and extending the electrophoresis time up to 3 h.

RESULTS

Design of *gag* HMA. Based on sequences available for HIV-1 group M subtypes A through H, primers were designed to amplify a fragment of approximately 460 bp covering a *gag* gene region coding for amino acid 132 of p24 to amino acid 20 of p7 (according to HIV-1 ELI; accession number KO3454). The experimental conditions were mainly as described for *env* HMA (ED31-ED33) (7), initially with a reference panel including 26 plasmids containing the *gag* gene of HIV-1 group M subtypes A to H. The choice of different subtype references was guided by their availability and by their phylogenetic classifications with respect to subtype representatives in distinguished subtype clusters for the *gag* HMA fragment analyzed (Table 1 and Fig. 1). Experimental conditions were obtained whereby reference isolates could be amplified by PCR and unambiguously subtyped by HMA. Altered experimental conditions relative to the gel composition (30% urea instead of 20%) allowed better distinction between *gag* subtype A, CRF AE_{CM240}, and CRF AG_{IbNG}.

Evaluation of *gag* HMA. An evaluation of *gag* HMA was done on a reference panel of 79 genetically confirmed HIV-1 group M isolates, isolated from individuals infected in different geographic regions, with the following *gag*/*env* subtypes: A/A, $n = 17$; A/D, $n = 2$; AG_{IbNG}, $n = 3$; A/G, $n = 2$; AE_{CM240}, $n = 7$; B/B, $n = 11$; B/A, $n = 1$; B/ND, $n = 2$; C/C, $n = 11$; D/D, $n = 7$; D/F, $n = 3$; F/F, $n = 7$; G/A, $n = 1$; G/G, $n = 1$; H/H, $n = 2$; and J/J, $n = 2$. For all of these isolates a positive PCR product was obtained. As indicated by phylogenetic analysis of the *gag* HMA references, two subtype D clusters were distinguished, which together with subtype B form one cluster (Fig. 1). Additional reference isolates were added to optimize subtype B (B/B, PIC63; B/B, PIC335) and D (D/D, VI761; D/F, PIC49) differentiation (Table 1). *gag* HMA was successful in determining the correct genetic subtype of 76 of 79 isolates (96%). Strains that could not be classified due to an intersubtype migration pattern belonged to subtype D ($n = 1$) and subtype J ($n = 2$). For the latter subtype no references were included.

TABLE 4. *gag*/*env* genetic subtyping of HIV-1 from plasma samples collected in Zambia, Cameroon, Kenya, and Bénin

Origin (total samples)[a]	Subtype[c]		n[b]
	gag	*env*[d]	
Zambia (61)	**C**	**C**	**60**
	A	A+C[d]	1 (dual infection)
Cameroon (66)	**A**	**A**	**21**
	D	**D**	**3**
	F	**F**	**1**
	G	**G**	**5**
	AG_{IbNG}	A	24
	AG_{IbNG}	F	1
	AG_{IbNG}	U	1
	A	D	1
	AE_{CM240}	AE_{CM240}	1
	A	AE_{CM240}	1
	A	F	1
	A	H	1
	G	A	3
	G	PCR−	1
	U	F	1
Kenya (86)	**A**	**A**	**39**
	C	**C**	**5**
	D	**D**	**17**
	A	C	2
	A	D	7
	A	G	1
	A	A+D	1 (dual infection)
	A	U	3
	C	D	2
	C	A	3
	D	A	5
	D+C	A+D	1 (multiple infection)
Bénin (39)	**A**	**A**	**15**
	G	**G**	**8**
	A	G	1
	AG_{IbNG}	A	14
	AG_{IbNG}	G	1

[a] Total of samples tested per country.
[b] n, number of samples per *gag*/*env* subtype.
[c] The nonrecombinant isolates (for the examined regions) are indicated in boldface, i.e., the same subtype in *gag* and *env*.
[d] A, subtype A and AG_{IbNG} could not be differentiated by *env* HMA.

Validation of *gag*/*env* HMA. *gag* HMA in combination with *env* HMA was validated on HIV-1-positive plasma samples from individuals who participated in the above-mentioned population-based survey.

A positive *gag* HMA PCR product was obtained for 242 of 252 (96%) of the analyzed plasma samples. The result obtained by *gag* HMA was confirmed by sequencing and phylogenetic analysis of the *gag* HMA fragment. For the 10 *gag* HMA fragment PCR-negative samples, an overlapping *gag* fragment was amplified. These samples were classified into subtypes A

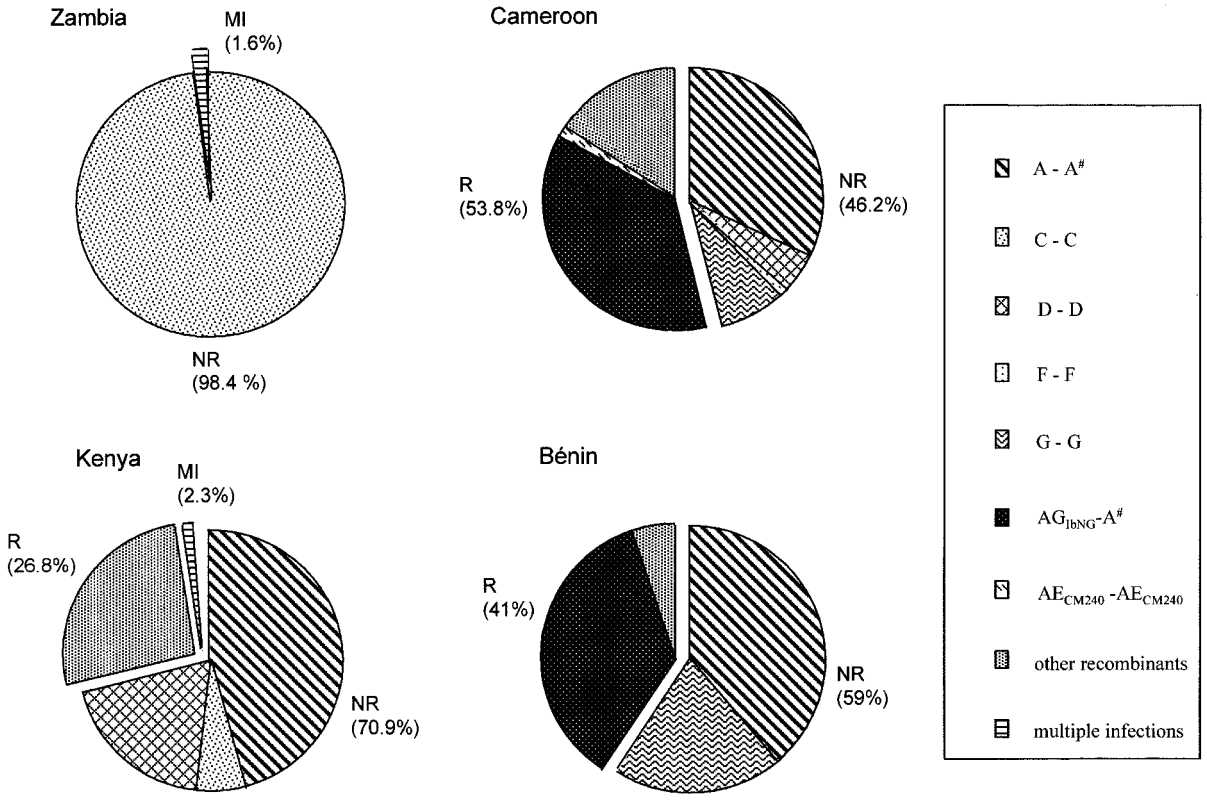

FIG. 2. Pie chart subtype representation for the four different countries based on *gag*/*env* regions. The pie is subdivided into pieces representing nonrecombinants (NR), recombinants (R), CRFs plus other recombinants, and multiple infections (MI). A, subtype A and AG_{IbNG} could not be differentiated by *env* HMA.

(CM, $n = 1$; KE, $n = 1$; and BJ, $n = 5$), C (KE, $n = 1$), and D (KE, $n = 2$) by sequencing and phylogenetic analysis of the *gag* HMA fragment. Of 242 samples, 6 could not be classified by *gag* HMA. These samples belonged to subtypes A (KE, $n = 1$), C (ZM, $n = 1$), D (KE, $n = 2$), dual D+C (KE, $n = 1$), and U (unclassified) (CM, $n = 1$, *env* subtype F) based on sequencing and phylogenetic analysis of the *gag* HMA fragments. *gag* subtype distribution results are presented in Table 2.

A positive PCR product for the *env* HMA ES7-ES8 fragment (6) was obtained for 238 of 252 (94.4%) of the analyzed plasma samples. For 13 of 14 PCR-negative samples a C2V3 coding *env* fragment could be amplified, sequenced, and phylogenetically analyzed. These samples were classified into subtypes A (CM, $n = 5$; KE, $n = 2$), D (KE, $n = 2$), AE_{CM240} (CM, $n = 1$), F (CM, $n = 1$), and U (KE, $n = 2$, *gag* subtype A) (Table 2). For the remaining sample no *env* fragment could be amplified (*gag* subtype G). Of 238 samples, 29 could not be unambiguously subtyped by *env* HMA. These samples were classified into subtypes A (CM, $n = 1$; KE, $n = 5$), C (KE, $n = 4$; ZM, $n = 1$), D (CM, $n = 4$; KE, $n = 3$), F (CM, $n = 1$), G (CM, $n = 3$; KE, $n = 1$), and U (ZM, $n = 1$, *gag* subtype C; CM, $n = 1$, *gag* subtype AG_{IbNG}; KE, $n = 1$, *gag* subtype A), as based on sequencing and phylogenetic analysis of the *env* C2V3 coding region. For three samples evidence of dual infection was obtained by cloning prior to subtyping by HMA and/or sequencing and phylogenetic analysis of the *env* C2V3 coding region: dual A+C (ZM, $n = 1$); dual A+D (KE, $n = 2$). *env* subtype distribution results are presented in Table 3.

The HIV-1 *gag* and *env* subtyping results from samples collected in Zambia, Cameroon, Kenya, and Bénin are presented in Table 4 and Fig. 2. Altogether, the frequency of HIV-1 intersubtype recombinants in this study was 53.8% (35 of 65) in Yaoundé, Cameroon; 41% (16 of 39) in Cotonou, Bénin; 26.8% (23 of 86) in Kisumu, Kenya; and 0% in Ndola, Zambia.

Differentiation between subtype A, CRF AE_{CM240}, and CRF AG_{IbNG}. CRFs AE_{CM240} and AG_{IbNG} each consist of variants that are genetically subtype A for the analyzed *gag* HMA fragment. Experimental *gag* HMA conditions as described for group M subtypes A to H to some extent also allow differentiation between subtype A, CRF AE_{CM240}, and CRF AG_{IbNG} (Fig. 3A). However this specific differentiation is much clearer when analyzed under altered gel conditions (30% urea) (Fig. 3B). In order to confirm correct classification of potential CRF AG_{IbNG} isolates, a *pol* gene fragment of 14 randomly chosen potential CRF AG_{IbNG} variants (CM, $n = 8$; BJ, $n = 6$) was PCR amplified, sequenced, and phylogenetically analyzed (Fig. 4). Clustering of these variants with CRF AG_{IbNG} representatives for this *pol* fragment supports the classification as CRF AG_{IbNG} variants (4).

DISCUSSION

The development of an efficacious vaccine against HIV-1 still remains one of the biggest challenges in the worldwide fight against HIV and AIDS. Major obstacles to vaccine development include the huge HIV variability and the fact that we do not know the correlates of protection in a vaccinated individual. The principle still holds that a vaccine should contain immunogenic characteristics of the prevalent HIV-1 subtype(s) circulating in a geographic region. We may need cocktails of different immunogens that are representative of each genotype and phenotype, in combination, or cocktails that are easily adapted to different geographic regions or over time. Up-to-date information on circulating HIV strains may thus be

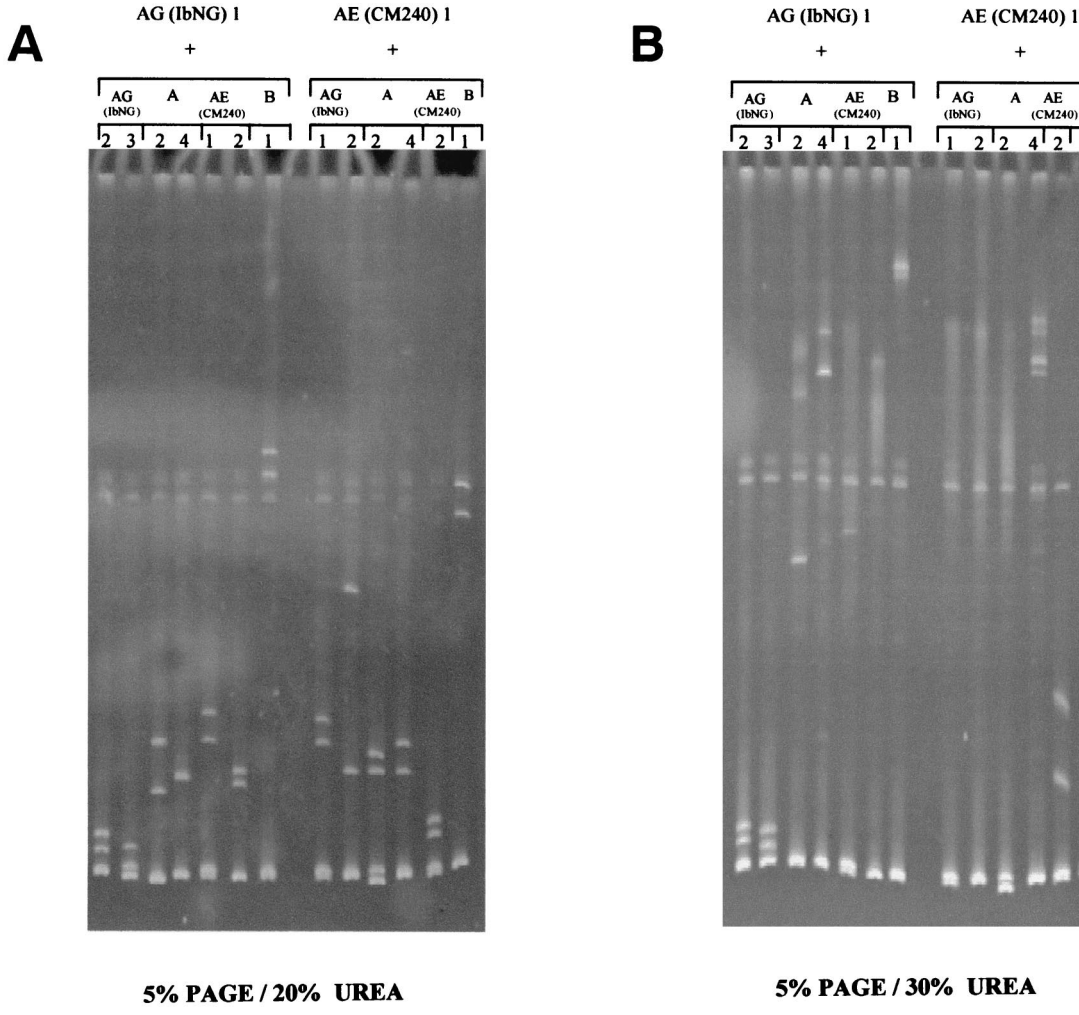

FIG. 3. HMAs of HIV-1 subtype A, AE_{CM240}, and AG_{IbNG} *gag* sequences on a 5% polyacrylamide gel containing 20% urea (A) and 30% urea (B). Isolate numbers are according to the references in Table 1.

of crucial importance. Although sequencing remains the most accurate approach for characterizing virus genomes, this method is time-consuming and requires a considerable investment in terms of equipment and reagents, as well as a lot of experience. This precludes the use of sequencing for monitoring the distribution of HIV strains in populations. *env* HMA has been evaluated as a reliable alternative subtyping method compared to sequencing and phylogenetic analysis (6, 13) and is sensitive, cost-effective, and applicable on a relatively large scale. It has already been successfully introduced by UNAIDS in several developing countries. Since intersubtype recombination is an important source of HIV-1 genetic variation, there was a need to extend HMA subtyping to cover two distinct HIV-1 regions: *gag* and *env*.

A reference panel was assembled based on available plasmids containing *gag* gene inserts and then evaluated with a panel of 79 genetically characterized samples. Two subtype B and two subtype D references were added to the reference panel in order to optimize differentiation among subtype B and D variants. The *gag* HMA subtype results correlated well with sequence and phylogenetic analysis, except for one subtype D isolate that was unclassified by *gag* HMA and two subtype J samples for which no representatives were included in the reference panel. Experimental conditions were found to distinguish between subtype A, CRF AE_{CM240}, and CRF AG_{IbNG} variants in subtype A, which was supported by sequencing and phylogenetic analysis of a *pol* gene fragment (4). The latter conditions only improve differentiation among subtype A, CRF AE_{CM240}, and CRF AG_{IbNG} and cannot be applied to subtyping other group M subtype strains. Subtyping by *gag* HMA correlated with sequencing and phylogenetic analysis of the same fragment.

The combination of *gag* and *env* subtyping results obtained on HIV-1-positive plasma samples from Bénin, Cameroon, Zambia, and Kenya revealed a high prevalence of a variety of intersubtype recombinants in Cameroon (53.8%), Kenya (26.8%), and Bénin (41%). These data indicate the relevance of subtyping two different gene fragments. Contrary to the lack of recombinants in Zambia, a high frequency of intersubtype recombinants documented in Cameroon, Kenya, and Bénin correlates with the distribution of multiple subtypes, as previously documented for Cameroon (25, 28), Kenya (24; E. M. Songok, H. Ichimura, P. M. Tukei, K. Kakimoto, P. Orege, N. Sakagami, and T. Kurimura, Abstr. 12th World AIDS Conf., abstr. 11188, 1998), and Bénin (12). AG_{IbNG}/F and AG_{IbNG}/G recombinants were identified in Cameroon and Bénin, respectively. The AG_{IbNG} recombinant subtype, as determined by *gag* HMA, is prevalent in Yaoundé, Cameroon (39.4%), and Co-

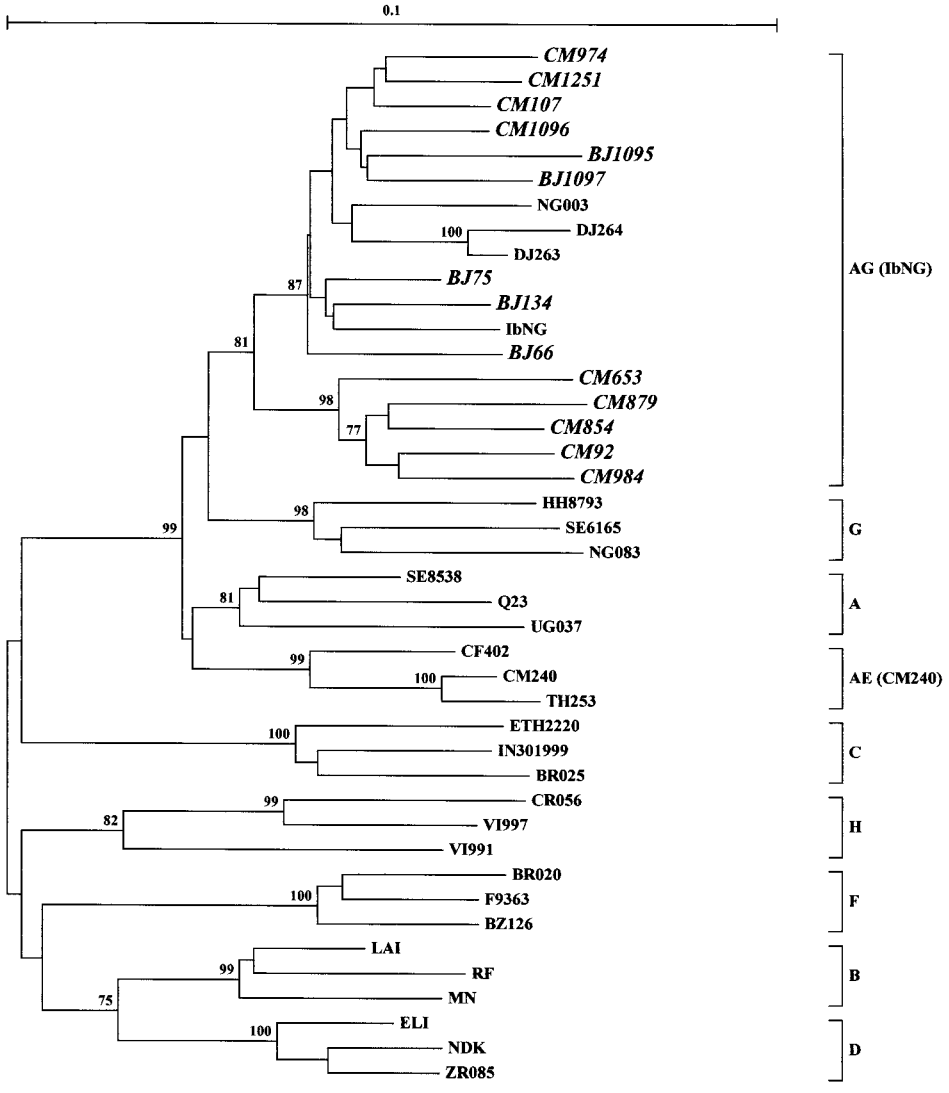

FIG. 4. Phylogenetic tree based on a *pol* gene region (750 bp; nucleotides 4377 to 5128; according to HIV-1 ELI) encoding integrase. Samples newly identified in this study are indicated in italic. A total of 2,000 bootstrap samples were analyzed. Bootstrap values are given in percentages at the internodes if they exceed the 70% level. The distance between two sequences is obtained by summing the lengths of the connecting horizontal branches by using the scale at the top. The tree is rooted arbitrarily.

tonou, Bénin (38.5%), and can therefore be considered as an epidemiologically important CRF. Reanalysis of previously documented *env* sequences encoding the region C2V3 to the start of gp41 (900 bp) of subtype A samples from Côte d'Ivoire (18) and Cameroon (25), in combination with *gag* HMA, indicates that CRF AG_{IbNG} strains were already prevalent in these countries (10 of 13 and 9 of 12, respectively) in the early 1990s. Furthermore, one isolate from Côte d'Ivoire was identified as an *env* AG_{IbNG}/*gag* A recombinant (data not shown). These results suggest that dual infection with different subtypes of HIV-1 may occur frequently in geographic regions where multiple HIV-1 subtypes cocirculate, giving rise to HIV-1 intersubtype recombinants that can be transmitted and spread in the population.

A different classification based on *gag* and *env* HMA obtained from the same sample is a strong indication for intersubtype recombination. However, no assumptions can be made regarding genome regions other than those analyzed in *gag* and *env* HMA. An isolate that is classified in the same subtype by *gag* and *env* HMA can still be an intersubtype recombinant. These results may thus have to be considered minimal estimates of the prevalence of recombinant strains. In the light of recent efforts to document full-length sequences of "pure" subtype and CRF references, the proposed reference panel is subject to improvement. Knowledge of subtype variants circulating in a particular geographic region may be used to add or replace subtype references for those variants most frequently encountered. The described *gag* HMA reference panel so far has allowed the subtyping of genetically confirmed HIV-1 group M *gag* subtype A to H strains and CRFs AE_{CM240} and AG_{IbNG} isolated from individuals infected in the following different geographic regions: subtype A and/or AG_{IbNG}, Côte d'Ivoire, Bénin, Kenya, Cameroon, Democratic Republic Congo, and Zambia; B, Belgium and the Phillipines; C, Kenya, Djibouti, Uganda, Democratic Republic Congo, Zambia, and Botswana; D, Kenya, Cameroon, and Democratic Republic Congo; AE_{CM240}, Cameroon and Thailand; F, Romania, Democratic Republic Congo, and Cameroon; G, Russia, Gabon, Bénin, and Cameroon; and H, Democratic Republic Congo and Gabon.

We believe that the development of a *gag* HMA makes an important contribution to subtyping strategies in surveillance,

in studies of biological characteristics of strains, and in surveys for the preparation of vaccine trials. Transfer of HIV-1 *gag*/*env* HMA to the developing countries may contribute to obtaining a more accurate estimate of the real prevalence of HIV-1 subtypes and intersubtype recombinants in those parts of the world where the technology for sequencing is not readily available. This will aid in the development and evaluation of HIV vaccines more suitable for use in developing countries.

ACKNOWLEDGMENTS

This work was supported by the Fonds voor Wetenschappelijk Onderzoek, Brussels, Belgium (grants 3-3301-96 and G.0134.97); the Flanders Interuniversity Institute for Biotechnology (VIB), Zwijnaarde, Belgium; the EC (grant IC18-CT97-02476); the Human Science Foundation (Tokyo, Japan); the United Kingdom Department for International Development (DFID, research project RD420); and UNAIDS. The work of the Study Group on Heterogeneity of HIV Epidemics in African Cities was supported by UNAIDS; the Agence Nationale de Recherche sur le SIDA/Ministère de la Coopération Française; DGXII of the EC (contract number ERBIC 18CT96-0109); SIDACTION (France); the Belgian Development Cooperation, Nairobi, Kenya; and the Fonds voor Wetenschappelijk Onderzoek (grant number K.V.E. 182, 1997), Brussels, Belgium. Linda Morison of the London School of Tropical Health and Medicine was supported by the British Medical Research Council.

REFERENCES

1. **Boom, R., C. J. A. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim-van Dillen, and J. van der Noordaa.** 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. **28:**495–503.
2. **Burke, D. S., and F. E. McCutchan.** 1997. Global distribution of human immunodeficiency virus-1 clades, p. 119–121. *In* V. T. De Vita, Jr., S. Hellman, and S. A. Rosenberg (ed.), AIDS: biology, diagnosis, treatment and prevention, 4th ed. Lippincott-Raven, Philadelphia, Pa.
3. **Carr, J. K., M. O. Salminen, C. Koch, D. Gotte, A. W. Artenstein, P. A. Hegerich, D. St. Louis, D. S. Burke, and F. E. McCutchan.** 1996. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 from Thailand. J. Virol. **70:**5935–5943.
4. **Carr, J. K., M. O. Salminen, J. Albert, E. Sanders-Buell, D. Gotte, D. L. Birx, and F. McCutchan.** 1998. Full genome sequences of human immunodeficiency virus type 1 subtypes G and A/G intersubtype recombinants. Virology **247:**22–31.
5. **Charneau, P., A. M. Borman, C. Quillant, D. Guétard, S. Chamaret, J. Cohen, G. Rémy, L. Montagnier, and F. Clavel.** 1994. Isolation and envelope sequence of a highly divergent HIV-1 isolate: definition of a new HIV-1 group. Virology **205:**247–253.
6. **Delwart, E. L., E. G. Sphaer, J. Louwagie, F. E. McCutchan, M. Grez, H. Rübsamen-Waigmann, and J. I. Mullins.** 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 genes. Science **262:**1257–1261.
7. **Delwart, E. L., M. P. Busch, M. L. Kalish, J. W. Mosley, and J. I. Mullins.** 1995. Rapid molecular epidemiology of human immunodeficiency virus transmission. AIDS Res. Hum. Retrovir. **11:**1081–1093.
8. **De Rijk, P., and R. De Wachter.** 1993. DCSE, an interactive tool for sequence alignment and secondary structure research. Comput. Appl. Biosci. **9:**735–740.
9. **Ellenberger, D. L., D. Pieniazek, J. Nkengasong, C.-C. Luo, S. Devare, C. Maurice, M. Janini, A. Ramos, C. Fridlund, D. J. Hu, I.-M. Coulibaly, E. Ekpini, S. Z. Wiktor, A. E. Greenberg, G. Schochetman, and M. A. Rayfield.** 1999. Genetic analysis of human immunodeficiency virus in Abidjan, Ivory Coast reveals predominance of HIV-1 subtype A and introduction of subtype G. AIDS Res. Hum. Retrovir. **15:**3–9.
10. **Fransen, K., P. Zhong, H. De Beenhouwer, G. Carpels, M. Peeters, J. Louwagie, W. Janssens, P. Piot, and G. van der Groen.** 1994. Design and evaluation of new, highly sensitive and specific primers for polymerase chain reaction detection of HIV-1 infected primary lymphocytes. Mol. Cell. Probes **8:**317–322. (Erratum, **9:**373, 1995.)
11. **Gao, F., D. L. Robertson, S. G. Morrison, H. W. Hui, S. Craig, J. Decker, P. N. Fultz, M. Girard, G. M. Shaw, B. H. Hahn, and P. M. Sharp.** 1996. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. J. Virol. **70:** 7013–7029.
12. **Heyndrickx, L., W. Janssens, M. Alary, K. Fransen, K. Vereecken, S. Coppens, B. Willems, N. Davo, A. Guèdèmè, E. Gabanizi, J. Joly, and G. van der Groen.** 1996. Genetic variability of HIV-1 in Bénin. AIDS Res. Hum. Retrovir. **12:**1495–1497.
13. **Heyndrickx, L., W. Janssens, S. Coppens, K. Vereecken, B. Willems, K. Fransen, R. Colebunders, M. Vandenbruaene, and G. van der Groen.** 1998. HIV-1 C2V3 *env* diversity among Belgian individuals. AIDS Res. Hum. Retrovir. **14:**1291–1296.
14. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz.** 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature **373:**123–126.
15. **Howard, T. M., D. O. Olaylele, and S. Rasheed.** 1994. Sequence analysis of the glycoprotein 120 coding region of a new HIV type 1 subtype A strain ($HIV\text{-}1_{IbNg}$) from Nigeria. AIDS Res. Hum. Retrovir. **10:**1755–1757.
16. **Hu, D. J., T. J. Dondero, M. A. Rayfield, J. R. George, G. Schochetman, H. W. Jaffe, C. C. Luo, M. L. Kalish, B. G. Weniger, C. P. Pau, C. A. Schable, and J. W. Curran.** 1996. The emerging genetic diversity of HIV—the importance of global surveillance for diagnostics, research, and prevention. J. Am. Med. Assoc. **275:**210–216.
17. **Janssens, W., A. Buvé, and J. Nkengasong.** 1997. The puzzle of HIV-1 subtypes in Africa. AIDS **11:**705–712.
18. **Janssens, W., L. Heyndrickx, Y. Van de Peer, A. Bouckaert, K. Fransen, J. Motte, G. M. Gershy-Damet, M. Peeters, P. Piot, and G. van der Groen.** 1994. Molecular phylogeny of part of the *env*-gene of HIV-1 isolated in Côte d'Ivoire. AIDS **8:**21–26.
19. **Korber, B., B. Hahn, B. Foley, J. W. Mellors, T. Leitner, G. Myers, F. McCutchan, and C. Kuiken.** 1997. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
20. **Louwagie, J., F. E. McCutchan, M. Peeters, T. P. Brennan, E. Sanders-Buell, G. A. Eddy, G. van der Groen, K. Fransen, G.-M. Gershy-Damet, R. Deleys, and D. S. Burke.** 1993. Phylogenetic analysis of *gag* genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. AIDS **7:** 769–780.
21. **Louwagie, J., W. Janssens, J. Mascola, L. Heyndrickx, P. Hegerich, G. van der Groen, F. E. McCutchan, G. Eddy, and D. Burke.** 1995. Genetic diversity of the HIV-1 envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. J. Virol. **69:**263–271.
22. **McCutchan, F. E., J. K. Carr, M. Bajani, E. Sanders-Buell, T. O. Harry, T. C. Stoeckli, K. E. Robbins, W. Gashua, A. Nasidi, W. Janssens, and M. L. Kalish.** 1999. Subtype G and multiple forms of A/G inter-subtype recombinant human immunodeficiency virus type 1 in Nigeria. Virology **254:**226–234.
23. **McCutchan, F. E., M. O. Salminen, J. K. Carr, and D. S. Burke.** 1996. HIV-1 genetic diversity. AIDS **10**(Suppl. 3)**:**S13–S20.
24. **Neilson, J. R., G. C. John, J. K. Carr, P. Lewis, J. K. Kreiss, S. Jackson, R. W. Nduati, D. Mbori-Ngacha, D. D. Panteleeff, S. Bodrug, C. Giachetti, M. A. Bott, B. A. Richardson, J. Bwayo, J. Ndinya-Achola, and J. Overbaugh.** 1999. Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya. J. Virol. **73:**4393–4403.
25. **Nkengasong, J. N., W. Janssens, L. Heyndrickx, K. Fransen, P. M. Ndumbe, J. Motte, A. Leonaers, M. Ngolle, P. Piot, and G. van der Groen.** 1994. Genotypic subtypes of HIV-1 in Cameroon. AIDS **8:**1405–1412.
26. **Novitsky, V. A., M. A. Montano, M. F. McLane, B. Renjifo, F. Vannberg, B. T. Foley, T. P. Ndung'u, M. Rahman, M. J. Makhema, R. Marlink, and M. Essex.** 1999. Molecular cloning and phylogenetic analysis of human immunodeficiency virus type 1 subtype C: a set of 23 full-length clones from Botswana. J. Virol. **73:**4427–4432.
27. **Simon, F., P. Mauclère, P. Roques, I. Loussert-Ajaka, M. C. Müller-Trutwin, S. Saragosti, M. C. Georges-Courbot, F. Barré-Sinoussi, and F. Brun-Vézinet.** 1998. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. Nat. Med. **4:**1032–1037.
28. **Takehisa, J., L. Zekeng, E. Ido, I. Mboudjeka, H. Moriyama, T. Miura, M. Yamashita, L. G. Gürtler, M. Hayami, and L. Kaptue.** 1998. Various types of HIV mixed infections in Cameroon. Virology **245:**1–10.
29. **Temin, H. M.** 1993. Retrovirus variation and reverse transcription: abnormal strand transfer result in retrovirus genetic variation. Proc. Natl. Acad. Sci. USA **90:**6900–6903.
30. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
31. **Van de Peer, Y., and R. De Wachter.** 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput. Appl. Biosci. **10:**569–570.