JOURNAL OF VIROLOGY, Sept. 1995, p. 5640–5649
0022-538X/95/$04.00+0
Copyright © 1995, American Society for Microbiology
Vol. 69, No. 9

Variability of Human Immunodeficiency Virus Type 1 Group O Strains Isolated from Cameroonian Patients Living in France

IBTISSAM LOUSSERT-AJAKA,[1,2] MARIE-LAURE CHAIX,[2] BETTE KORBER,[3,4] FRANCK LETOURNEUR,[2] EMMANUEL GOMAS,[2] ETHAN ALLEN,[3,4] THOAI-DONG LY,[5] FRANÇOISE BRUN-VÉZINET,[1] FRANÇOIS SIMON,[1] AND SENTOB SARAGOSTI[2]*

Laboratoire de Virologie, Hôpital Bichat-Claude Bernard, 75877 Paris Cedex 18,[1] Institut Cochin de Genetique Moleculaire, Institut National de la Santé et de la Recherche Médicale U363, Hôpital Cochin,[2] and Institut Fournier,[5] 75014 Paris, France; and Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico 87545[3]; and the Santa Fe Institute, Santa Fe, New Mexico 87501[4]

Received 4 January 1995/Accepted 1 June 1995

Human immunodeficiency virus type 1 (HIV-1) nucleotide sequences encoding p24[Gag] and the Env C2V3 region were obtained from seven patients who were selected on the basis of having paradoxical seronegativity on a subset of HIV enzyme-linked immunosorbent assay detection kits and having atypical Western blot (immunoblot) reactivity. Sequence analyses showed that all of these strains were more closely related to the recently described Cameroonian HIV isolates of group O (HIV-1 outlier) than to group M (HIV-1 major). All seven patients had Cameroonian origins but were living in France at the time the blood samples were taken. Characterization of a large number of group M strains has to date revealed eight distinct genetic subtypes (A to H). Genetic distances between sequences from available group O isolates were generally comparable to those observed in M intersubtype sequence comparisons, showing that the group O viruses are genetically very diverse. Analysis of sequences from these seven new viral strains, combined with the three previously characterized group O strains, revealed few discernable phylogenetic clustering patterns among the 10 patients' viral sequences. The level of diversity among group O sequences suggests that they may have a comparable (or greater) age than the M group sequences, although for unknown reasons, the latter group dispersed first and is the dominant lineage in the pandemic.

Two types of human immunodeficiency viruses (HIV-1 and -2) have been described as the etiologic agents of AIDS (2, 8, 16, 32). These two HIV types are genetically distinct, notably with respect to the regulatory genes *vpu* and *vpx*, which are unique to HIV-1 and HIV-2, respectively. Numerous HIV-1 strains have been isolated from diverse geographic locations worldwide, and nucleic acid sequence analyses have identified at least eight distinct HIV-1 clades, designated subtypes A to H (23, 35, 37, 42, 46). Each of these subtypes is approximately equidistant from the others in a phylogenetic pattern that has been described as a star phylogeny (28). No correlation between HIV-1 subtypes and viral pathogenicity or transmissibility has yet been found, although subtype-specific mutational patterns in the V3 region have been observed (28). HIV-2 has recently been shown to present a greater genetic and biological diversity than anticipated (17). In the case of HIV-2, a correlation between ease of viral isolation and genetic subtype has been noted (17, 56), and HIV-2 may have a decreased pathogenic potential relative to HIV-1 (40).

In 1990, a preliminary sequence of a Cameroonian HIV-1 isolate (ANT70) was described as being radically different from other HIV sequences (10). A similar virus (MVP5180) was isolated from a second Cameroonian patient, and it was found to cluster phylogenetically with ANT70 (18). These viruses (18, 63) have been fully sequenced and were classified as HIV-1 in view of their overall genomic structure. Because of their great genetic distances from other HIV-1 isolates, they were originally considered outliers and so were shown together in sequence alignments with a virus isolated from a chimpanzee captured in the wild in Gabon (CPZGAB) (21). DNA sequence analyses showed ANT70 and MVP5180 had on average 65 to 70% similarity to HIV-1 and 56% similarity to HIV-2. Recently, Charneau et al. suggested splitting the HIV-1 viruses into two groups, group M (for major) and group O (for outlier), containing the sequences ANT70, MVP5180, and VAU (7). (VAU was isolated from a Caucasian woman living in France with no known ties to Africa.)

By competitive immunoblotting with the MVP5180 group O strain, an estimated prevalence of HIV-1 group O-infected individuals of 7 to 8% was found in Yaoundé, Cameroon (18). This estimated prevalence was confirmed by a retrospective study with an ANT70 V3 synthetic peptide-based enzyme-linked immunosorbent assay (ELISA) to test sera collected from HIV-seropositive patients in Cameroon and Gabon (24, 49). This prevalence could be an underestimate, however, because the HIV-1 V3 loop region used for the peptide ELISA is a highly variable region among group O viruses, as illustrated in this article. The most recent survey, in contrast, estimated the prevalence of HIV-1 group O infection to be 2% or less, accounting for less than 10% of HIV-1-infected patients in Cameroon (65). In a previous report, we described the characterization of nine patients living in France (eight from Cameroon) who are infected with group O viruses (34). These HIV variants present a public health challenge, since several of them escape detection by certain conventional screening tests (58). In addition, the serum samples of most of these patients

* Corresponding author. Mailing address: Institut Cochin de Genetique Moleculaire, INSERM U363, Hôpital Cochin, 22 rue Méchain, 75014 Paris, France. Phone: 33-1-40.78.26.45 or 33-1-46.33.37.75. Fax: 33-1-46.33.92.97 or 33-1-45.89.74.05. Electronic mail address: sentob@citi2.fr.

TABLE 1. Clinical and virological status of the seven HIV-1 group O-infected patients

Patient	Age (yr)	Sex[a]	CD4 count (10^6/liter)	CDC classification	Plasma culture	Coculture
BCF01	44	M	91	IVA	+	+
BCF02	38	F	19	IVC1	ND[b]	+
BCF03	45	M	97	IVC1	+	+
BCF06	22	M	10	IVC1	ND	+
BCF07	29	F	384	II	ND	+
BCF08	22	F	420	II	−	+
BCF11	68	F	575	II	ND	+

[a] M, male; F, female.
[b] ND, not determined.

displayed atypical and incomplete Western blot (immunoblot) reactivity (34, 58). One of the nine patients of this group was lost to follow-up; another patient was analyzed in a recent study by Charneau et al. (7).

Here, we report the genetic characterization of viral sequences from the seven remaining patients. Sequences of two nonoverlapping genomic regions encoding the p24[Gag] capsid protein and the C2V3 region of Env were obtained and analyzed. The sequences were based on those of strains isolated by coculture. The results of phylogenetic analyses indicated that the group O viruses are all very distant from HIV-1 group M subtypes as well as from the CPZ virus. Our results confirmed that these HIV-1 viruses have a distinct lineage compared with HIV-1 group M. The two groups, M and O, form a phylogenetic pattern that roughly corresponds to a two-star phylogeny.

MATERIALS AND METHODS

Patients. The seven Cameroonian patients included in this study were monitored in Parisian hospitals. They were selected for their HIV serological profile, atypical Western blot, and reactivity to a synthetic peptide assay based on the MVP5180 strain; the serological assays were conducted in three different laboratories in Paris (34). As summarized in Table 1, four patients were classified as CDC IV and three patients were classified as CDC II at the time of virus isolation (according to the Centers for Disease Control and Prevention [CDC] classification). Their ages varied between 22 and 44 years for the symptomatic patients and between 22 and 68 years for the asymptomatic patients. The CD4 counts for the symptomatic patients were $<100 \times 10^6$/liter, and the CD4 counts for the asymptomatic patients were $>380 \times 10^6$/liter. For each patient, the number following the letters BCF corresponds to the number of the serum sample described by Simon et al. (58).

Cultures. Whole blood from the seven patients was collected on heparin lithium. Peripheral blood mononuclear cells (PBMCs) were isolated in a Ficoll-Hypaque (Pharmacia) density gradient. The cells were washed twice with RPMI 1640 and resuspended (2.10^6/ml) in culture medium supplemented with 20% fetal calf serum, 10% human interleukin-2, 5% L-glutamine, and antibiotics (150 μg of streptomycin per ml, 250 U of penicillin G per ml). Patient cells (10^6) were cocultured with an equal number of phytohemagglutinin-stimulated PBMCs obtained from healthy donors (duplicates in 24-well plates [Costar]). Medium was replaced twice weekly, and 3×10^5 fresh cells were added at days 7, 14, and 21. Viral replication was monitored in culture supernatants for 28 days by reverse transcriptase assay (micromethod) and by a polyclonal anti-p24 detection kit (ELAVIA p24 Ag; Sanofi-Diagnostic Pasteur). High-reverse transcriptase-positive supernatants were frozen at −80°C for further studies.

PCR, cloning, and sequencing. DNA from cocultures and fresh PBMCs from infected patients were extracted with phenol-chloroform, precipitated with ethanol, and quantified spectrophotometrically. PCRs were carried out in 100-μl reaction mixtures containing 10 mmol of Tris-HCl per liter (pH 8.3), 50 mmol of KCl per liter, 2 mmol of $MgCl_2$ per liter, 0.2 mmol (each) deoxynucleoside triphosphate (dNTP) per liter, 40 pmol of each primer, 2.5 U of *Taq* polymerase (Perkin-Elmer Cetus, St. Quentin Yvelines, France), and 1 μg of cell DNA. The primers used were Env5-Cam [5′-GACTCGAGT(R)GTTACTTGTACACATG GCAT-3′] and Env3-Cam (5′-GAGAATTCACAATAAAAGAACTCCATGA CAGT-3′) for the *env* region and Gag5-Cam (5′-GACTCGAGCAGGGACAA ATGGTACATCA-3′) and Gag3-Cam (5′-GAGAATTCAGTAGCTTGCTCA GCTCTTAAT-3′) for the *gag* region, corresponding to nucleotides 6991 to 7012, 7421 to 7396, 1250 to 1269, and 1768 to 1747 of the HIV-1 ANT70 sequence, respectively, and selected from conserved regions in the sequences of HIV-1 subtype O isolates ANT70 and MVP5180. Samples were submitted to 40 amplification cycles, each consisting of three successive steps: denaturation at 94°C for 1 min, annealing of primers at 50°C for *gag* and 55°C for *env* for 1 min, and extension at 72°C for 1 min. In the first cycle, denaturation was performed for 4 min, and in the last cycle, denaturation was extended over 5 min. Amplified products were digested with *Xho*I and *Eco*RI restriction enzymes, purified, and cloned in an M13mp18 vector digested by *Sal*I and *Eco*RI.

Sequencing. For each patient, between three and six clones were sequenced with the PRISM Ready Reaction Ampli*Taq* dye primer −21M13 kit (Perkin-Elmer) on an automated DNA sequencer (Applied Biosystems 373A sequencer), but only the consensus sequence derived from the sets of cloned sequences was retained for further analysis. The amino acid sequence for each patient was translated from the nucleotide consensus sequence. The sequences from the viral cultures tended to be highly similar, and the consensus for each intrastrain viral sequence set was defined by picking the most common nucleotide found in each position. Because this study was directed towards exploring group O viral diversity between infected patients and not within a patient, a single consensus sequence per individual was used for the study rather than the complete set of sequences.

Phylogenetic analysis. DNA sequences were aligned with the multiple aligned sequence editor MASE (13). Tree topologies were inferred by neighbor joining with the Kimura two-parameter distance matrix (PHYLIP) (14), with a transition/transversion ratio of 1.3. Weighted parsimony (PAUP and MacClade) (20, 39, 62) was also used; character state changes were weighted as previously described (28). Confidence values for individual branches of the resulting trees were evaluated by a bootstrap analysis in which 100 bootstrap trees were generated with PHYLIP and PAUP for neighbor joining and parsimony trees, respectively. Only unambiguous positions in the alignments were included, leaving 255 nucleotides with 198 varied sites for the C2V3 region trees and 396 nucleotides with 195 varied sites for the p24 region trees.

The sequences used in the construction of the *env* C2V3 phylogenetic tree correspond to the following strains: JRCSF (50); HU08448 and Z321 (61); DJ258A, DJ258A, and SE365A (38); NOF (11); NDK (60); TN243 (41); VI525A, DJ259A, and SE365A (38); NOF (11); NDK (60); TN243 (41); VI525A and LBV217 (23); BZ126A and BZ163A (38); CARELO (40a); and CPZGAB (21). The sequences used in the construction of the p24[*gag*] phylogenetic trees corresponded to the following strains: DJ258, VI313, DJ259, VI525, VI557, UG280, UG270, UG274, VI174, BZ162, LBV217, VI191, and CI4 (37); JRCSF (50); and CPZGAB (21).

Synonymous versus nonsynonymous substitutions in the *env* C2V3 region. The proportion of observed synonymous substitutions divided by the potential number of synonymous substitutions (*ps*) and the proportion of observed nonsynonymous substitutions divided by the number of possible nonsynonymous substitutions (*pn*) were calculated by the method of Nei and Gojobori (47). Variable domains with insertions and deletions were excluded from the analysis. The calculations of slopes were based on *ps* values of less than 0.3 to minimize problems due to homoplasy.

Nucleotide sequence accession numbers. The full set of *env* and *gag* sequences is available through GenBank. The accession numbers for *env* are as follows: BCF01, U24566; BCF02, U24562; BCF03, U24567; BCF06, U24568; BCF07, U24563; BCF08, U24564; and BCF11, U24565. Those for *gag* are as follows: BCF01B, U24706; BCF02B, U24707; BCF03B, U24708; BCF06B, U24711; BCF07B, U24709; BCF08B, U24710; and BCF11B, U24712.

RESULTS

Virus isolation. In the seven patients' PBMCs, virus was detectable between days 5 and 21 of coculture, via monitoring of reverse transcriptase activity and p24 antigen reactivity. The optical density observed for the p24 antigenemia was weak compared with that for classical HIV-1, as already described (7, 18), with values resembling those obtained with HIV-2 (59). This result reflects the poor reactivity against group O p24[Gag] of the antibodies generated against an HIV-1 group M p24 antigen used in this assay. Characteristic retroviral cytopathic effect with syncytium formation was observed between days 4 and 15 for all viruses. Virus isolation from plasma was performed with three patients and was successful for only two symptomatic patients (CDC IVA and CDC IVC1) (Table 1). The negative patient was asymptomatic and had CD4 cell counts greater than 400×10^6/liter (Table 1).

Nucleic acid sequence. PCR amplifications with HIV-1 group M- and HIV-2-specific primers for the *env*, *pol*, and *nef* genes were negative (1, 34). In two cases, with HIV-1 group M *gag*-specific primers (SK38 and SK39), we were able to obtain an amplified product. Analysis of the sequence of the amplified

TABLE 2. Intra- and intergroup genetic distance

Virus type	% Genetic distance					
	gag			*env* C2V3		
	Group O	CPZGAB	Group M	Group O	CPZGAB	Group M
Group M	24.2–31.6	22.0–32.0	2.5–23.9	38.9–49.0	39.0–47.7	2.8–22.7
CPZGAB	27.0–29.0			44.4–51.5		
Group O	4.3–14.3			4.6–30.3		

DNA allowed us to characterize these isolates as belonging to group O. Primers derived from the consensus sequences of ANT70 and MVP5180 (Env5-Cam, Env3-Cam, Gag5-Cam, and Gag3-Cam) allowed amplification all of the isolates analyzed in this study. The PCR amplification and sequencing was done on DNA extracted from cultured PBMCs, although for five patients, direct PCR was also positive on DNA extracted from uncultured PBMCs. A 497-bp fragment within the p24*gag* coding region and a 405-bp fragment spanning the C2V3 region of *env* were amplified. By using these primers, direct PCR amplification was successful for all group O isolates tested to date. The specificity of these primers was tested on samples obtained from patients infected with HIV-1 group M viruses. Neither group M strains isolated from French patients (n = 34) nor those isolated from African patients (n = 24) could be amplified.

***gag* sequences.** The genetic distance between group O and group M strains and CPZGAB virus were calculated. The intragroup Hamming distances ranged from 4.03 to 14.3% for group O and from 2.5 to 23.9% for group M. The distances between group O sequences versus CPZGAB were 27 to 29%, while the distances between CPZGAB and group M sequences ranged from 22 to 32% (Table 2). These are conservative estimates, because only regions which could be unambiguously aligned were retained for the comparisons. We attempted to select a set of group M viral sequences that were sampled within approximately the same time frame as the group O samples discussed here, but otherwise the samples were arbitrarily selected to represent each of the group M subtypes (A to H) (46). The same sets of nucleic acid sequences included in the similarity comparisons were used for the phylogenetic analysis. While the numbers presented here are useful for the sake of comparison, there are sampling limitations for both the M and group O sequence sets. Therefore, there would certainly be some variation in the intra- and intergroup values if other characteristic isolates were chosen. The amino acid sequence alignments are shown in Fig. 1. Among the 133 amino acids (aa) analyzed, 25 were variable in group O sequences. In 13 positions, amino acids were replaced with amino acids of the same character in terms of biochemical similarity or charge. A comparison of translated p24Gag group O sequences revealed that the synonymous nucleotide substitutions with no effect on protein sequences were far more common than nonsynonymous nucleotide substitutions in this region, analogous to what is found in HIV-1 and HIV-2 (27).

***env* C2V3 region.** For the seven patients, 318 to 323 bp of sequence spanning the C2V3 region was characterized. The intergroup sequence similarity between group O isolates and group M subtypes was calculated, again with the same set of sequences for the similarity comparisons and for the generation of phylogenetic trees. The *env* nucleotide sequence divergence among the seven Cameroonian isolates ranged from 4.6 to 30.3%, while the divergence among clades A to H ranged from 2.8 to 22.7% (Table 2). Interestingly, the divergence between group O isolates and the CPZGAB isolate is higher than the divergence between group M isolates and CPZGAB (44.4 to 51.5 and 39 to 47.7%, respectively).

Deduced amino acid sequences of the *env* C2V3 region. The amino acid sequence alignments are shown in Fig. 2. In contrast to what was found for group M strains (6, 29, 46), the variability inside the V3 loop is higher than the variability outside the V3 loop among group O strains. Among the 62 aa characterized upstream and downstream of the V3 loop, 32 were conserved. In most group M strains, the V3 loop is composed of 34 or 35 aa, while the V3 loop of group O isolates is composed of 36 (ANT70 and VAU), 37 (BCF02, BCF01, BCF03, BCF07, and BCF11), 38 (BCF06 and BCF08), or 39 (MVP5180) aa. Only 9 aa, including the two cysteines, were absolutely conserved among all group O strains. Some intrapatient variability was observed; however, the small number of clones sequenced for each patient and the possible emergence of selected isolates yielded by coculture, as observed by Murphy et al. (44), prevent us from drawing any conclusions. Except for the GP motif (shown in boldface), the octameric sequence at the tip of the loop of the 10 isolates (including

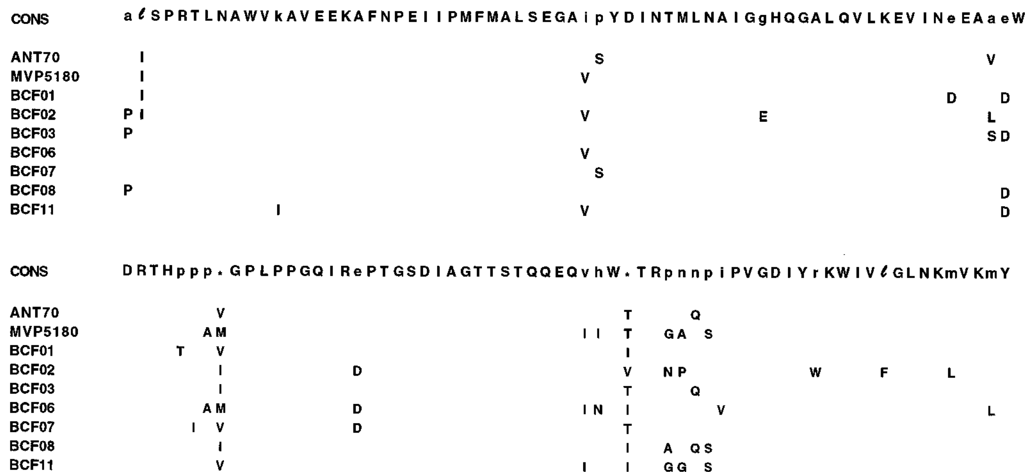

FIG. 1. Alignment of deduced amino acid sequences of a fragment of the p24*gag* of group O Cameroonian strains. The top line corresponds to the consensus (CONS) sequence. Asterisks indicate positions which contain fewer than five identical amino acids.

C2 Region

```
CONS      CTHGIkPtVSTqLIINGTIS•gkIrimgknIsdsgkNIIvTLNttI--NmT--

ANT70          R         N    K    M A D  EG        NS L  N
MVP5180                  N    RE       N TE A       N P   N
VAU/HAM                  N    K N T N  N     E  LI  N --  NI IA
BCF01                    N    EKG      N  KT E      NVS   NI
BCF02                    N    E    M A N     Q      N     N
BCF03            A       N    K      A N TNT N      NS    NI
BCF06                   MN    R       RN T NT       N S   N
BCF07                    N    K    L A N     Q      N     N
BCF08                H   N  I E E      N RENA       NS    N
BCF11            A       N    E       QN            NK VNMNI
```

V3 Loop

```
CONS      C•Rp-gn••vQel•iGPmawySmgℓ--a•n••••sR•AyC •Yn•t•W•••

ANT70      E   -QIDI  MR          I  GGTAGNS  A    K NA D G
MVP5180    I -E IAE  D YT   R R  T KRSNNTSPR  V    T NK V E
VAU/HAM    E    NQTI K MA        A   -SNTKGDT A    N SA D N
BCF01      H    NLS   MK   LS         ANSSIK  V    N ST E T
BCF02      Q    HQT    R              NGNGSE  R    E NT N I
BCF03      N    RGIK - G    SV  GS  ADLGGNNN  I    D DI K NET
BCF06      M -K RGKI R AT  LR V  AA KTESQNTG  I    M NN E I
BCF07      H    NLK    K          I  EAENIPD  K    D NA K V
BCF08      E  E NLTI   HS       L    KRNTTVR  S H  K NT N E
BCF11      T -D DQK    G   LS    SI   EDSAKNT A    N SASS K
```

FIG. 2. Alignment of deduced amino acid sequences of the *env* C2V3 region of group O Cameroonian strains. Only group O sequences were aligned. In addition to the sequences of the seven patients we characterized, the previously published sequences were added (ANT70, MVP5180, and VAU). Dots indicate gaps introduced to improve the alignment, and asterisks indicate ambiguity in the consensus sequence. The presence of a potential N-linked glycosylation site is shown by a boldface asparagine [N-X-(T or S)].

ANT70, MVP5180, and VAU) is very heterogeneous: RI**GP**MAWY (n = 2), GI**GP**LSWY, GI**GP**MSVY, KI**GP**MAWY, KI**GP**LSWY, YT**GP**MRWR, AT**GP**LRWV, MA**GP**MAWY, and GI**GP**LSWY.

Synonymous and nonsynonymous. The group M and group O strains used for the sequence distance matrices and tree analysis were also used for synonymous and nonsynonymous pairwise comparisons (Fig. 3). Silent substitutions were essentially saturated in comparisons of group M and group O sequences from either gene. The relationship between the nonsynonymous and synonymous substitutions in the p24 regions of viruses from both groups displayed similar slopes in plots of pn versus ps. This is illustrated in Fig. 3A, where the slopes indicating the pn/ps ratio are 0.15 for group O comparisons (Fig. 3A.A) and 0.13 for group M comparisons (Fig. 3A.B). This indicates that the synonymous substitutions are approximately 6.6-fold more frequent than the nonsynonymous substitutions among the group O strains and are 7.7-fold more frequent than nonsynonymous substitutions among the group M viruses; the difference between the two groups was not statistically significant. The silent substitutions between group O and group M sequences are completely saturated (Fig. 3A.C).

In the C2V3 region, however, the group O viruses revealed a significantly greater variation at the protein level, with pn/ps slopes of 0.58 for group O viruses (Fig. 3B.A) and 0.45 for group M viruses (Fig. 3B.B). The significance was tested by a Monte Carlo resampling of points with a random-no-replacement method and calculating the differences in the slopes for 100 randomized data sets; in the 100 randomizations, no differences in slope greater than 0.05 were observed. In the real data, the difference in slopes is 0.13. The intragroup distribution of the ps values for both the M and O groups in the C2V3 region and in p24 are comparable, suggesting that the samples representing these groups may have had similar lengths of time to diverge.

Phylogenetic analysis of the group O isolates and characteristic group M strains. Phylogenetic trees were constructed from the group O sequences available: 10 sequences for C2V3 and 9 sequences for *gag*. In addition, group M representative strains were also included. In both trees, the sequences characteristic of each group M subtype clustered together, indicating that the regions we chose are informative for subtype determination.

Figure 4 shows a weighted parsimony tree with C2V3 (Fig. 4A) and p24 region (Fig. 4B) sequences. The numbers that appear at the nodes are bootstrap values calculated with nonweighted parsimony. The numbers represent the frequency with which the node recurred out of 100 bootstrap replicates and are restricted to values over 50 of 100. Neighbor-joining trees gave comparable bootstrap proportions (data not shown).

It is apparent that the HIV-1 viruses we analyzed should be divided into two groups: group M, as proposed by Charneau et al. (7), and group O. There are only two O clusters that are reproducibly detected in both regions. The first is the clustering of sequences BCF02 and BCF07 (Fig. 4A), although the association is less clear in *gag* than it is in the C2V3 region. Another potentially interesting association is that between MVP5180, BCF11, and BCF06 in *gag*. These sequences have a highly conserved region in the $p24^{gag}$ gene, but the relationship between BCF11 and the other two sequences in the C2V3 region is lost, suggesting the possibility of recombination. Recombination was looked for by signature analysis (25) in the regions available, but other than the possibility noted in BCF11, no evidence for recombination was found.

N glycosylation. N glycosylation may play an important role in infection. The glycosylation pattern may influence infectivity (31), the immunogenicity of the V3 loop (3), and protein folding (33). Overall, the Env protein of HIV-1 group M contains more than 20 potential glycosylation sites (33); the C2V3 region we analyzed contains seven potential N-linked glycosylation sites (glycosylation sites are represented with a boldface N in the sequences shown in Fig. 2). In the C2 region, the first and fourth sites are absolutely conserved (10 of 10), while the two other glycosylation sites are well conserved (8 of 10). Site

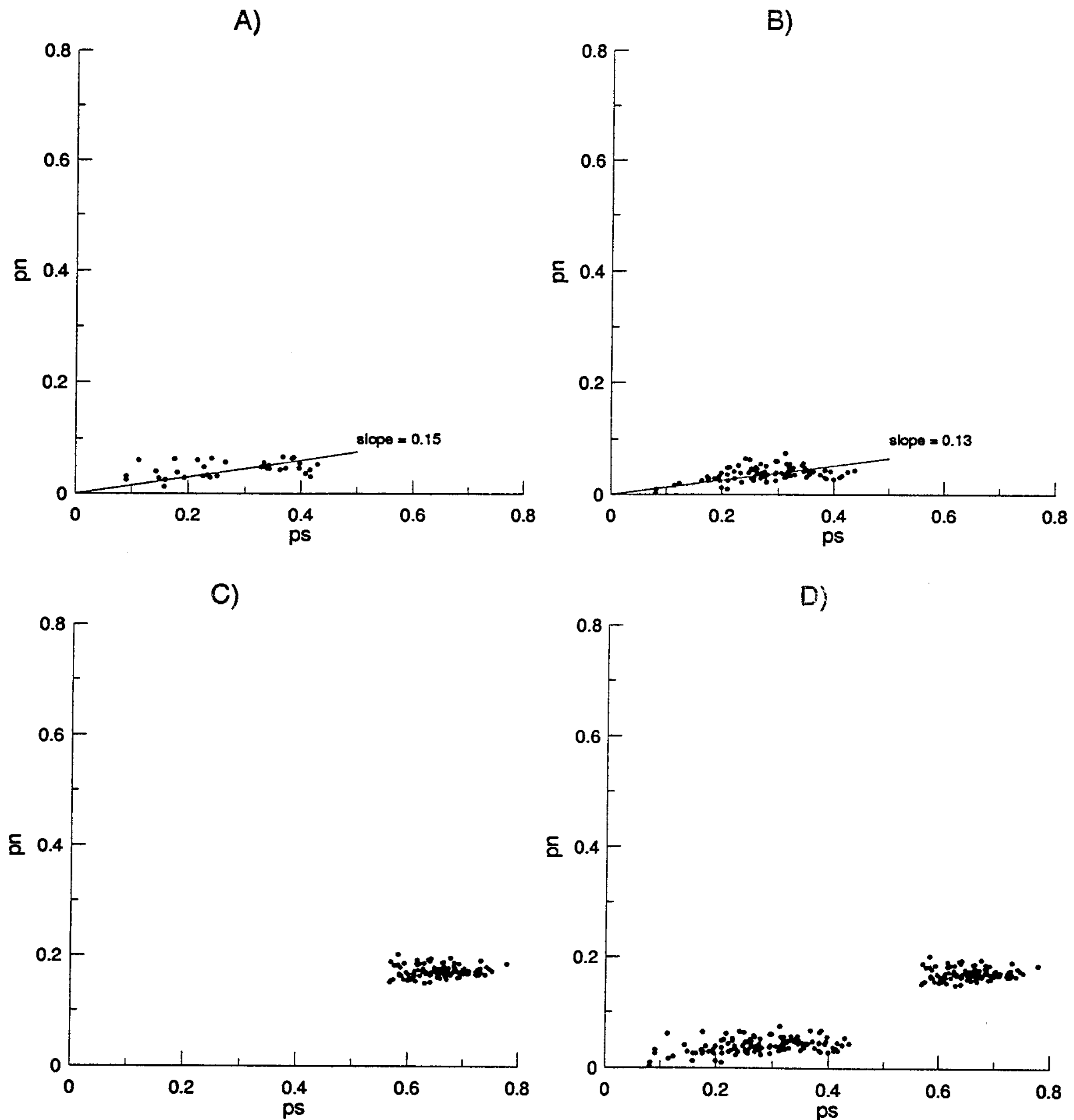

FIG. 3. Ratio of synonymous to nonsynonymous substitutions in the region encoding CAp24[Gag] (A) and C2V3. The sequences used are divided into two groups: M and O. Pairwise comparisons of all sequences included in this study were made, synonymous/nonsynonymous (ps/pn) values were determined, and the *ps* and *pn* for each pair of sequences were calculated. The *ps* values are plotted on the abscissa, and the *pn* values are plotted on the ordinate, so the slope is an indication of the *pn*/*ps* ratio. The slopes were calculated only for points with *ps* values of less than 0.3 to minimize problems due to homoplasy and were calculated through the origin. Panels C and D are included to indicate that the sequence divergence is at saturation for synonymous substitutions between group M and O sequences. Panels A, group O sequence *ps*/*pn* values; panels B, group M sequence *ps*/*pn* values; panels C, group M versus O sequence intergroup *ps*/*pn* values; panels D, group M and group O sequence intergroup and intragroup *ps*/*pn* values. Only intergroup values are shown for panels D.

5, spanning the first cysteine of the V3 loop, is conserved in most group M isolates from throughout the world (5, 12, 19, 22, 30, 41, 44, 48, 51, 53, 54, 67) with the exception of clade C isolates (46). This site is missing in group O isolates as well as in HIV-2 isolates (4, 45) (Fig. 5). The site next to the cysteine was moved forward and was present only in three group O isolates. In the carboxyl-terminal part of the V3 loop, a highly variable region among Cameroonian group O strains, a glyco-

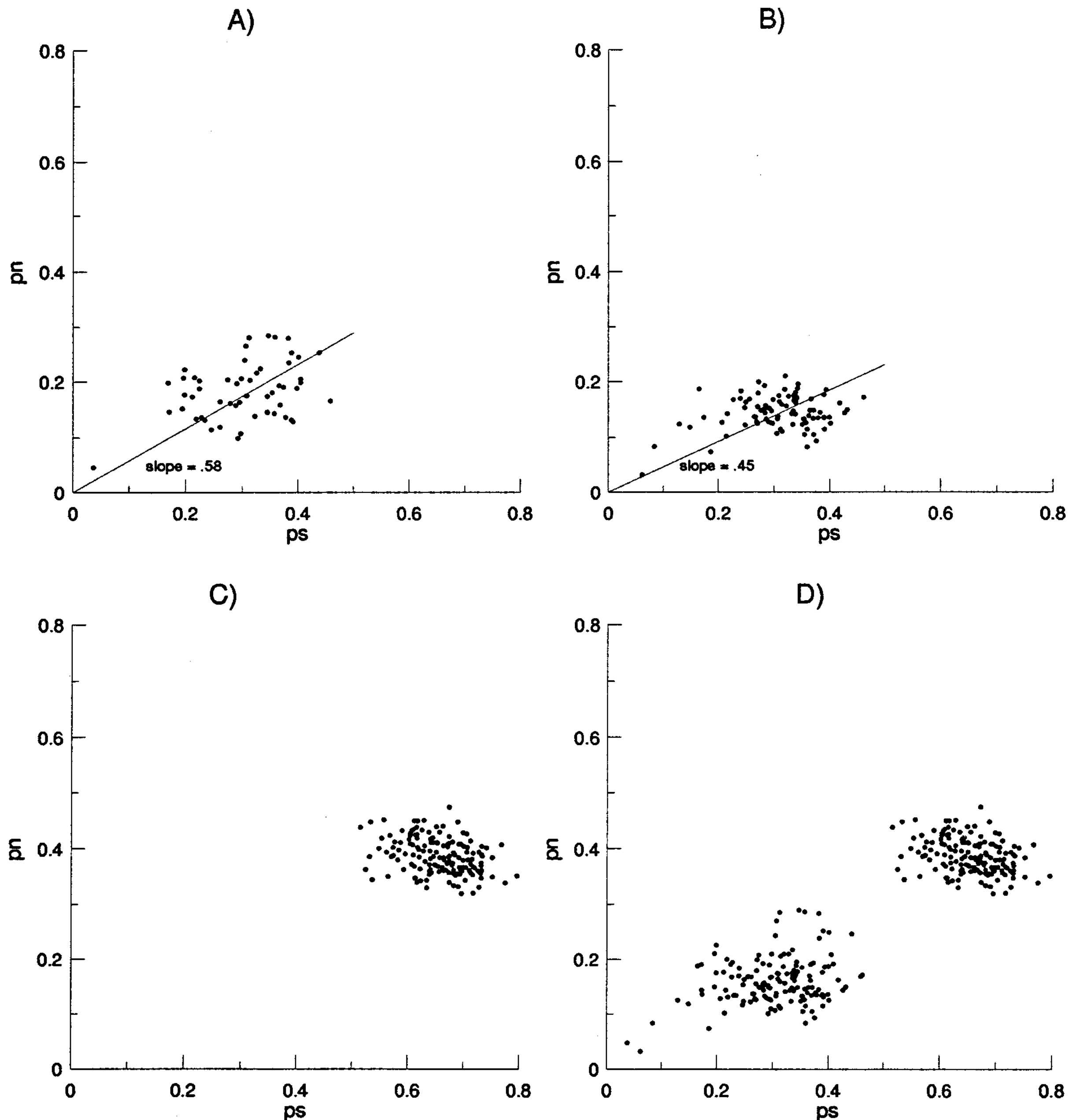

FIG. 3—*Continued.*

sylation site was present in 6 of the 10 isolates. In the MVP5180 isolate, two adjacent sites were found. Interestingly three strains (BCF06, BCF07, and BCF11) do not have potential N-glycosylation sites in the V3 loop (for the BCF07 isolate, a site was found in one of four sequences in the C terminus). Downstream of the V3 loop, all strains had a glycosylation site but at variable positions, similar to subtype B strains isolated from French patients (6). The potential and consensus sites present in HIV-1 and HIV-2 are represented in Fig. 5.

DISCUSSION

The two full-length sequences of ANT70 and MVP5180 and the *env* sequence of VAU represented the most divergent HIV-1 isolates previously reported. It was clear that these viruses were antigenically and genetically distinct from the well-characterized group M (major) virus, so they became classified as a new group of HIV-1 viruses, group O. Because of the high level of divergence between HIV-1 group M and group O

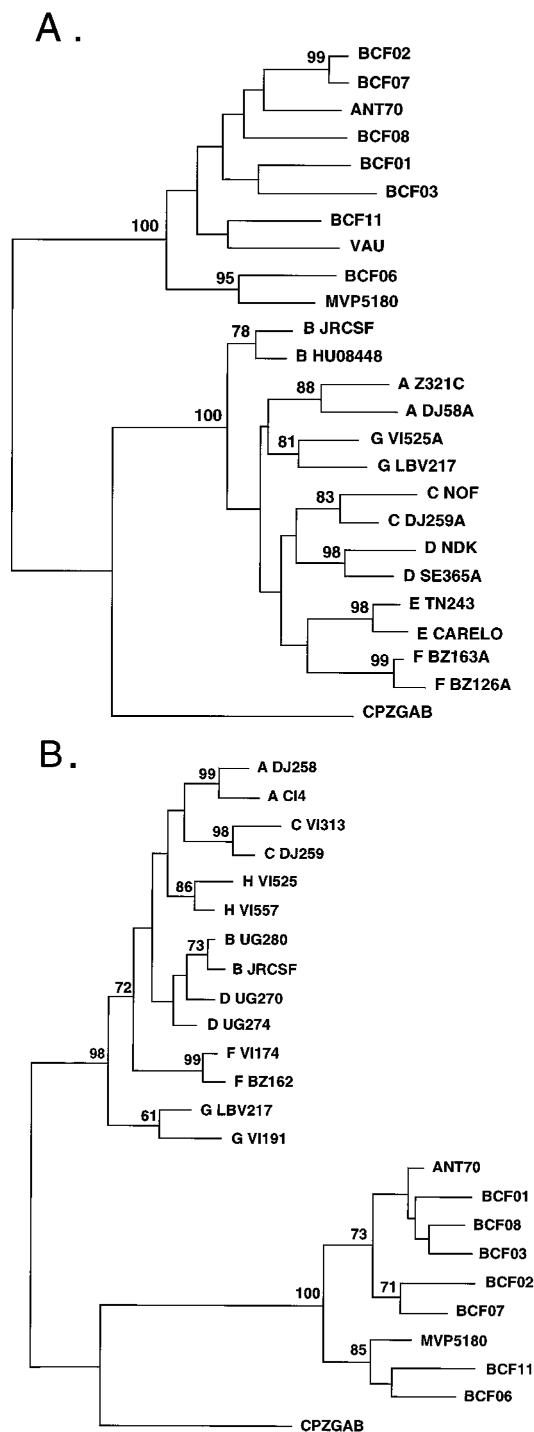

FIG. 4. Phylogenetic tree analysis comparing the HIV-1 *env* C2V3 region and CAp24^{gag} coding sequences. A weighted parsimony tree is shown. The numbers given at the branch points are the 50% threshold majority consensus values for 100 bootstrap replicates. The lengths of the horizontal branches are proportional and do not reflect single base changes because weighted parsimony was used, but should be indicative of relative evolutionary distances; vertical distances are for clarity only. These trees were based on alignments encoding nucleotide sequences from which columns containing gaps have been deleted. (A) *env* C2V3 region. (B) CAp24^{gag}.

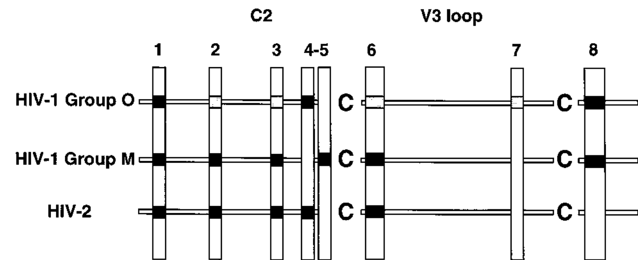

FIG. 5. Potential N-linked glycosylation sites of HIV-1 and HIV-2 strains. The figure represents a schematic view of the consensus N-glycosylation sites in HIV-1 group O and group M and HIV-2. The potential sites present in all strains are represented as black boxes, and potential sites present occasionally are represented as open boxes.

isolates, one could propose naming group O HIV-3 instead. However, the group O strains have been defined as HIV-1 viruses for three reasons: (i) a *vpu*-like gene characteristic of HIV-1 is present, (ii) the *env* and *nef* genes are not overlapping as they are in HIV-2 genomes, and (iii) the TAR sequence is similar in terms of structure and size to HIV-1 TAR (10). Moreover, the relative ease of virus isolation from cells and plasma indicates that in culture, HIV-1 group O viruses behave more similarly to HIV-1 group M isolates than to HIV-2 isolates (59).

The identification of multiple group O-infected individuals in France indicates that the spread of viruses belonging to this group has already started in Europe. Viruses of group O are likely to be transmitted vertically (5a) and horizontally through heterosexual routes, leading to an ever wider distribution, as has been seen for HIV-1 group M viruses and for HIV-2 (9). The pathogenicity of these viruses is under investigation, but a comprehensive study of their natural history will require several more years. Nevertheless, it is already clear that these viruses can be associated with AIDS, since two infected patients have already died of AIDS-like syndromes (MVP5180 and VAU/HAM) and four of our patients reached CDC stage IV. All of the PCR and serological data indicated that these patients were not coinfected with HIV-1 group M subtypes or with human T-cell leukemia virus type 1 or 2.

Analysis of *gag* sequences confirmed the discrimination of HIV-1 strains into group M and group O strains. Most of the substitutions observed in the fragment of *gag* we studied were synonymous mutations, with no effect on protein sequences. Only 25 of 133 positions were nonsynonymous substitutions among the group O sequences, and in most cases, the variations were conservative in nature.

The V3 region of gp120 encompasses a neutralizing antibody epitope (52, 55, 66). The loop contains B- and T-cell epitopes and is a determinant of infectivity, membrane fusion, and syncytium formation (15, 43). Comparison of the V3 loops of group O strains showed striking divergence in terms of size and amino acid composition. As was previously shown for the ANT70, MVP5180, and VAU/HAM isolates, the tip of the loop of the seven group O strains we characterized did not contain the GPG motif, which is present in most group M subtypes. The two motifs found were GPM ($n = 7$) and GPL ($n = 3$). There is also a great rate of protein divergence in the C2V3 region, as seen by comparing *pn*/*ps* ratios. This may simply be due to a sampling artifact or alternatively it may have important implications. Shpaer and Mullins have observed a correlation between high rates of nonsynonymous mutations and the pathogenic potential of different primate lentiviruses (57); the nonsynonymous rate is even higher in the group O

C2V3 region sequences tested than observed in group M viruses. Also, the remarkable variation in this region may result in a greater capacity to elude Env-based detection methods among group O viruses than among group M viruses.

Analysis of the potential N-linked glycosylations shows a pattern different from that found in most group M strains. The most striking difference concerns site 5, adjacent to the first cysteine (NCT), which is absent in the 10 group O strains and in group M subtype C (46). The most N-terminal site (site 6) present in the V3 loop in all group M clades is present only in three group O strains, and an additional site (site 7) is present in the C terminus of the loop in 6 of 10 strains. An alteration in the position of sites 7 and 8 was also observed. Three strains do not contain any sites within the V3 loop (Fig. 5). The design of a vaccine inducing neutralizing antibodies directed against the V3 loop is a difficult task that can be impaired by the modification of the antigenicity of the loop by posttranslational N glycosylation. The divergent sequences observed, the presence of new motifs at the tip of the V3 loop, and the absence in three patients or the presence in six patients of new N-glycosylation sites suggest that a vaccine directed against HIV-1 group M strains may be ineffective against group O strains because of variation in glycosylation patterns as well as overall protein variation.

Analysis of *gag* sequences confirmed the discrimination of HIV-1 strains into group M and group O strains. Most of the substitutions observed in the fragment of *gag* we studied were synonymous mutations with no effect on protein sequences. Only 25 of 133 positions were nonsynonymous substitutions, and in most cases, the variations were conservative. The relative conservation of $p24^{Gag}$ is also observed in group M strains.

Since the characterization of the first HIV isolate (64), the accumulation of sequences of a number of different HIV-1 isolates has shown that this virus can be divided into distinct subtypes. In group M, eight subtypes, A to H, have already been characterized, primarily on the basis of *env* and *gag* sequence comparisons. The different group M subtypes have been found to be roughly equidistant from one another, suggesting that at some point in time, founder populations of the subtypes separated from a common pool. If one assumes that the different subtypes are evolving at a similar pace, then the subtypes may have originated in the same evolutionary time frame, moving out to form distinct viral lineages because of some unknown epidemiological factor. Similar phylogenetic analysis made with the 10 group O strains showed limited phylogenetic clustering, with only two pairs of sequences grouping into two clades in both the p24 and C2V3 regions. There are two scenarios that are consistent with this observation. The first is that the lack of a clear branching structure in the group O trees could be an artifact of small sample size, and clear phylogenetic associations among group O viruses may eventually emerge as more sequences become available, revealing an underlying subtype pattern similar to that of group M viruses. The second is that the evolutionary and epidemiological history of the group O and M viruses may have been different: in the case of M group viruses, yielding the distinctive lineages corresponding to the subtypes, and in the case of group O viruses, simply radiating out from an original ancestral virus. On the basis of the intersubtype genetic distances found among group M viruses and the similar distances found among group O viruses, it is interesting to speculate that the two groups, M and O, form a double-star phylogenetic pattern. The similar range of *ps* values observed within the two groups (M and O) indicates they may have started spreading in their human hosts in roughly comparable time frames.

The low prevalence and recent emergence of this HIV-1 group raise the question of its origin. Low prevalence is often correlated with low infectivity, low pathogenicity, or recent introduction. However, the first evidence indicates that the group O viruses do not have reduced pathogenicity compared with group M viruses, and the genetic analysis suggests that group O may be comparable to group M in terms of age. Thus, a socioeconomic hypothesis (epidemiologic factors) may account for the delayed emergence of group O compared with group M. We cannot rule out the hypothesis that group O viruses simply have been detected as HIV-1 and are already more widespread than we currently imagine.

Nucleotide sequences of HIV-1 obtained from African patients display great diversity (26, 36, 37), with Western Central Africa being, to our knowledge, the region where the most divergent group M strains have been observed (44, 49). The appearance of HIV-1 group O in this geographic region (24, 48) suggests that it may be a region of great importance for continued studies of viral variation in both humans and other primate species. In this region, not only are most of the HIV-1 group M subtypes and HIV-1 group O present, but HIV-2 and human T-cell leukemia viruses type 1 and 2 are present as well (9a). The extraordinary potential for variation in HIV-1 must be taken into account in HIV diagnosis and in vaccine and drug development. Better understanding of the extent of variation of HIV-1 both provides insight into HIV's evolutionary history and helps us understand its potential for variation in the future.

ACKNOWLEDGMENTS

We thank Jean Paul Levy and Gerry Myers for helpful discussions. We also thank Kersti MacInnes from the Database and Analysis staff of the Los Alamos National Laboratory.

This work was supported by the A.N.R.S. (French National AIDS Research Agency) and by grants from INSERM. I.L.-A. and M.-L.C. were supported by A.N.R.S. fellowships. B.K. and E.A. were supported by the AIDS Division of the National Institute of Allergy and Infectious disease. B.K. is also supported by the Pediatric AIDS Foundation.

REFERENCES

1. **Agut, H., D. Candotti, B. Rabanel, J.-M. Huraux, G. Remy, D. Ingrand, T. Tabary, C. Chippaux, S. Chamaret, D. Guetard, C. Dauguet, and L. Montagnier.** 1992. Isolation of atypical HIV-1-related retrovirus from AIDS patient. Lancet **340:**681–682.
2. **Barré-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Axler-Blin, F. Brun-Vézinet, C. Rouzioux, W. Rozenbaum, and L. Montagnier.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science **220:**868–871.
3. **Benjouad, A., J.-C. Gluckman, H. Rochat, L. Montagnier, and E. Bahraoui.** 1992. Influence of carbohydrate moieties on the immunogenicity of human immunodeficiency virus type 1 recombinant gp160. J. Virol. **66:**2473–2483.
4. **Boeri, E., A. Giri, F. Lillo, G. Ferrari, O. E. Varnier, A. Ferro, S. Sabbatani, W. C. Saxinger, and G. Franchini.** 1992. In vivo genetic variability of the human immunodeficiency virus type 2 V3 region. J. Virol. **66:**4546–4550.
5. **Buonaguro, L., D. Greco, M. L. Tornesello, M. Rago, M. Zaccarelli, F. Curcio, G. Castello, N. Sguazzo, G. Pipolo, E. Beth-Giraldo, F. M. Buonaguro, and G. Giraldo.** 1994. Analysis of HIV-1 env gene V3 loop sequence in a southern Italian cohort of intravenous drug users. AIDS **8:**268–269.

5a. **Chaix, M. L., et al.** Unpublished data.

6. **Chaix, M. L., C. Chappey, I. Couillin, W. Rozenbaum, J. P. Levy, and S. Saragosti.** 1993. Diversity of the V3 region of HIV in Paris, France. AIDS **7:**1199–1204.
7. **Charneau, P., A. M. Borman, C. Quillent, 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.
8. **Clavel, F., D. Guétard, F. Brun-Vézinet, S. Chamaret, M. A. Rey, M. O. Santos-Ferreira, A. G. Laurent, C. Dauguet, C. Katlama, C. Rouzioux, D. Klatzmann, J. L. Champalimaud, and L. Montagnier.** 1986. Isolation of a new human retrovirus from West African patients with AIDS. Science **233:**343–346.
9. **De Cock, K. M., G. Adjorlolo, E. Ekpini, T. Sibailly, J. Kouadio, M. Maran,**

K. Brattegaard, K. M. Vetter, R. Doorly, and H. D. Gayle. 1993. Epidemiology and transmission of HIV-2: why there is no HIV-2 pandemic. JAMA **270:**2083–2086.

9a. **Delaporte, E., N. Monplaisir, J. Louwagie, M. Peeters, Y. Martin-Prével, J.-P. Louis, A. Trebucq, L. Bedjabaga, S. Ossari, C. Honoré, B. Larouzé, L. d'Auriol, G. van der Groen, and P. Piot.** 1991. Prevalence of HTLV-I and HTLV-II infection in Gabon, Africa: comparison of the serological and PCR results. Int. J. Cancer **49:**373–376.

10. **De Leys, R., B. Vanderborght, M. Vanden Haesevelde, L. Heyndrickx, A. van Geel, C. Wauters, R. Bernaerts, E. Saman, P. Nijs, B. Willems, H. Taelman, G. van der Groen, P. Piot, T. Tersmette, J. G. Huisman, and H. Van Heuverswyn.** 1991. Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of West-Central African origin. J. Virol. **64:**1207–1216.

11. **Dietrich, U., M. Grez, H. von Briesen, B. Panhans, M. Geibendorfer, H. Kuhnel, J. Maniar, G. Mahambre, W. B. Becker, M. L. B. Becker, and H. Rubsamen-Waigmann.** 1993. HIV-1 strains from India are highly divergent from prototypic African and US/European strains, but are linked to a South African isolate. AIDS **7:**23–27.

12. **Dumitrescu, O., M. L. Kalish, S. C. Kliks, C. I. Bandea, and J. A. Levy.** 1994. Characterisation of human immunodeficiency virus type 1 isolates from children in Romania: identification of a new envelope subtype. J. Infect. Dis. **169:**281–288.

13. **Faulkner, D. V., and A. Jurka.** 1988. Multiple aligned sequence editor (MASE). Trends Biochem. Sci. **13:**321–322.

14. **Felsenstein, J.** 1993. PHYLIP (phylogeny interference package) version 3.5c. Department of Genetics, University of Washington, Seattle.

15. **Freed, E. O., D. J. Meyers, and R. Risser.** 1991. Identification of the principal neutralizing determinant of human immunodeficiency virus type 1 as a fusion domain. J. Virol. **65:**190–194.

16. **Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Palker, R. Redfield, J. Oleske, B. Safai, G. White, P. Foster, and P. D. Markham.** 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science **224:**500–503.

17. **Gao, F., L. Yue, D. L. Robertson, S. C. Hill, H. Hui, R. J. Biggar, A. E. Neequaye, T. M. Whelan, D. D. Ho, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1994. Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. J. Virol. **68:**7433–7447.

18. **Gürtler, L. G., P. H. Hauser, J. Eberle, A. von Brunn, S. Knapp, L. Zenkeng, J. M. Tsague, and L. Kaptue.** 1994. A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon. J. Virol. **68:**1581–1585.

19. **Hattori, T., K. Shiozaki, Y. Eda, S. Tokiyoshi, S. Matsushita, H. Inaba, M. Fujimaki, T. Meguro, K. Yamada, M. Honda, K. Nishikawa, and K. Takatsuki.** 1991. Characteristics of the principal neutralizing determinant of HIV-1 prevalent in Japan. AIDS Res. Hum. Retroviruses **7:**825–830.

20. **Hillis, D. M., J. P. Huelsenbeck, and C. V. Cunningham.** 1994. Application and accuracy of molecular phylogenies. Science **264:**671–677.

21. **Huet, T., R. Cheynier, A. Meyerhans, G. Roelants, and S. Wain-Hobson.** 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature (London) **345:**356–359.

22. **Ichimura, H., S. C. Kliks, S. Visrutaratna, C.-Y. Ou, M. L. Kalish, and J. A. Levy.** 1994. Biological, serological, and genetic characterisation of HIV-1 subtype E isolates from northern Thailand. AIDS Res. Hum. Retroviruses **10:**263–269.

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

24. **Janssens, W., J. N. Nkengasong, L. Heyndrickx, K. Fransen, P. M. Ndumbe, E. Delaporte, M. Peeters, J.-L. Perret, A. Ndoumou, C. Atende, P. Piot, and G. Van Der Groen.** 1994. Further evidence of the presence of genetically aberrant HIV-1 strains in Cameroon and Gabon. AIDS **8:**1012–1013.

25. **Korber, B., and G. Myers.** 1992. Signature pattern analysis: a method for assessing viral sequence relatedness. AIDS Res. Hum. Retroviruses **8:**1549–1560.

26. **Korber, B., G. Myers, and S. Wolinsky.** 1992. Analysis of inter- and intrapatient variation in the V3 loop of the HIV-1 envelope protein. Vaccines **9:**75–79.

27. **Korber, B. T. M., A. Allen, A. D. Farmer, and G. Myers.** Heterogeneity of HIV-1 and HIV-2. AIDS: 1995. A year in review, in press.

28. **Korber, B. T. M., K. MacInnes, R. F. Smith, and G. Myers.** 1994. Mutational trends in V3 loop protein sequences observed in different genetic lineages of human immunodeficiency virus type 1. J. Virol. **68:**6730–6744.

29. **Kuiken, C. L., G. Zwart, E. Baan, R. A. Coutinho, J. A. R. van den Hoek, and J. Goudsmit.** 1993. Increasing antigenicity and genetic diversity of the HIV-1 V3 domain in the course of the AIDS epidemic. Proc. Natl. Acad. Sci. USA **90:**9061–9065.

30. **Larosa, G. J., J. P. Davide, K. Weinhod, J. A. Waterbury, A. T. Profy, J. A. Lewis, A. J. Langlois, G. R. Dreesman, R. N. Boswell, P. Shadduck, L. H. Holley, M. Karplus, D. P. Bolognesi, T. J. Matthews, E. A. Emini, and S. D. Putney.** 1990. Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant. Science **249:**932–935.

31. **Lee, W. R., W. Syu, B. Du, M. Matsuda, S. Tan, A. Wolf, M. Essex, and T. Lee.** 1992. Non-random distribution of gp 120 N-linked glycosylation sites important for infectivity of human immunodeficiency viruses type 1. Proc. Natl. Acad. Sci. USA **89:**2213–2217.

32. **Levy, J. A., A. D. Hoffman, S. M. Kramer, J. A. Landis, J. M. Shimabukuro, and L. S. Oshiro.** 1984. Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. Science **225:**840–842.

33. **Li, Y., L. Luo, N. Rasool, and C. Y. Kang.** 1993. Glycosylation is necessary for the correct folding of human immunodeficiency virus gp120 in CD4 binding. J. Virol. **67:**584–588.

34. **Loussert-Ajaka, I., T. D. Ly, M. L. Chaix, D. Ingrand, S. Saragosti, A. M. Couroucé, F. Brun-Vézinet, and F. Simon.** 1994. HIV-1/HIV-2 seronegativity in HIV-1 subtype O infected patients. Lancet **343:**1393–1394.

35. **Louwagie, J., W. Janssens, L. J. Mascola, L. Heyndrickx, P. Hegerich, G. van der Groen, F. E. McCutchan, and D. S. Burke.** 1995. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. J. Virol. **69:**263–271.

36. **Louwagie, J., F. McCutchan, J. Mascola, G. Eddy, K. Fransen, M. Peeters, G. Van Der Groen, and D. Burke.** 1993. Session II. Technical challenges and scientific issues. Genetic subtypes of HIV-1. AIDS Res. Hum. Retroviruses **9**(Suppl. 1)**:**S147–S150.

37. **Louwagie, J., F. E. McCutchan, M. Peeters, T. P. Brennan, E. Scanders-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.

38. **Louwagie, J. J., E. Delwart, J. I. Mullins, F. E. McCutchan, G. Eddy, and D. S. Burke.** Unpublished data.

39. **Maddison, W. P., and D. R. Maddison.** 1992. MacClade: analysis of phylogeny and character evolution. Sinauer Associates, Inc. Sunderland, Mass.

40. **Marlink, R., P. Kanki, I. Thior, K. Travers, G. Eisen, T. Siby, I. Traore, C. C. Hsieh, M. C. Dia, E. H. Gueye, J. Hellinger, A. Gueye-Ndiaye, J. L. Sankale, I. Ndoye, S. Mboup, and J. L. Sankale.** 1994. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. Science **265:**1587–1590.

40a. **Mathiot, C., et al.** Unpublished data.

41. **McCutchan, F. E., P. Hegerich, T. P. Brennan, P. Phanuphak, P. Singharaj, A. Jugsudee, P. W. Berman, A. M. Gray, A. K. Fowler, and D. S. Burke.** 1992. Genetic variants of HIV-1 in Thailand. AIDS Res. Hum. Retroviruses **8:**1887–1895.

42. **McCutchan, F. E., B. L. P. Ungar, P. Hegerich, C. R. Roberts, A. K. Fowler, S. K. Hora, P. L. Perine, and D. S. Burke.** 1992. Genetic analysis for HIV-1 isolates from Zambia and an expanded phylogenetic tree for HIV-1. AIDS **5:**441–449.

43. **Moore, J. P., and P. L. Nara.** 1991. The role of the V3 loop of gp120 in HIV infection. AIDS **5**(Suppl. 2)**:**S21–S33.

44. **Murphy, E., B. Korber, M. C. Georges-Courbot, B. You, A. Pinter, D. Cook, M. P. Kieny, A. Georges, C. Mathiot, F. Barre-Sinoussi, and M. Girard.** 1993. Diversity of V3 region sequences of human immunodeficiency viruses type 1 from the Central African Republic. AIDS Res. Hum. Retroviruses **9:**997–1006.

45. **Myers, G., B. Korber, J. A. Berzofsky, R. F. Smith, and G. N. Pavlakis.** 1992. Human retroviruses and AIDS 1992: a compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Theoretical Biology and Biophysics (T10), Los Alamos, N. Mex.

46. **Myers, G., B. Korber, S. Wain-Hobson, R. F. Smith, and G. N. Pavlakis.** 1993. Human retroviruses and AIDS 1993: a compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Theoretical Biology and Biophysics (T10), Los Alamos, N. Mex.

47. **Nei, M., and T. Gojobori.** 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3:**418–426.

48. **Nkengasong, J. N., W. Janssens, L. Heyndrickx, K. Fransen, P. M. Ndumbe, J. Motte, A. Leonaers, M. Ngolle, J. Ayuk, P. Piot, and G. Van der Groen.** 1994. Genotypic subtypes of HIV-1 in Cameroon. AIDS **8:**1405–1412.

49. **Nkengasong, J. N., M. Peeters, M. Vanden Haesevelde, S. S. Musi, B. Willems, P. M. Ndumbe, E. Delaporte, J. L. Perret, P. Piot, and G. Van den Groen.** 1993. Antigenic evidence of the presence of the aberrant HIV-1 ANT70 virus in Cameroon and Gabon. AIDS **7:**1536–1538.

50. **O'Brien, W. A., Y. Koyanagi, A. Namazie, J. Q. Zhao, A. Diagne, K. Idler, and I. S. Y. Chen.** 1990. HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 which do not include the CD4 binding domain. Nature (London) **348:**69–73.

51. **Oram, J. D., R. G. Downing, M. Roff, N. Serwankambo, J. C. S. Clegg, A. S. R. Featherstone, and J. C. Booth.** 1991. Sequence analysis of the V3 loop regions of the env genes of Ugandan human immunodeficiency proviruses. AIDS Res. Hum. Retroviruses **7:**605–614.

52. **Palker, T. J., M. E. Clark, A. J. Langlois, T. J. Matthews, K. J. Weinhold, R. R. Randall, D. P. Bolognesi, and B. F. Haynes.** 1988. Type-specific neutralization of the human immunodeficiency virus with antibodies to Env-

encoded synthetic peptides. Proc. Natl. Acad. Sci. USA **85:**1932–1936.

53. **Potts, K. E., M. L. Kalish, C. I. Bandea, G. M. Orloff, M. St. Louis, C. Brown, N. Malanda, M. Kavuka, G. Schochetman, C. Y. Ou, and W. L. Heyward.** 1993. Genetic diversity of human immunodeficiency virus type 1 strains in Kinshasa, Zaire. AIDS Res. Hum. Retroviruses **9:**613–618.
54. **Potts, K. E., M. L. Kalish, T. Lott, G. Orloff, C. C. Luo, M. A. Bernard, C. Brites Alves, R. Badaro, J. Suleiman, O. Ferreira, G. Schochetman, W. D. Johnson, C. Y. Ou, J. L. Ho, and the Brazilian Collaborative AIDS Research Group.** 1993. Genetic heterogeneity of the V3 region of the HIV-1 envelope glycoprotein in Brazil. AIDS **7:**1191–1197.
55. **Rusche, J. R., K. Javaherian, C. McDanal, J. Petro, D. L. Lynn, R. Grimalia, A. Langlois, R. C. Gallo, L. O. Arthur, P. J. Fischinger, D. P. Bolognesi, S. D. Putney, and T. J. Matthews.** 1988. Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24 amino-acid sequence of the viral envelope, gp 120. Proc. Natl. Acad. Sci. USA **85:**3198–3202.
56. **Sharp, P. M., D. L. Robertson, F. Gao, and B. Hahn.** 1994. Origins and diversity of human immunodeficiency viruses. A year in review. AIDS **8:**S27–S43.
57. **Shpaer, E. G., and J. I. Mullins.** 1993. Rates of amino acid change in the envelope protein correlate with pathogenicity of primate lentiviruses. J. Mol. Evol. **37:**57–65.
58. **Simon, F., T. D. Ly, A. Baillou-Beaufils, V. Schneider-Fauveau, J. de Saint-Martin, I. Loussert-Ajaka, M. L. Chaix, S. Saragosti, A. M. Courroucé, D. Ingrand, C. Janot, and F. Brun-Vézinet.** 1994. Sensitivity of screening kits for anti-HIV-1 subtype O antibodies. AIDS **8:**1628–1629.
59. **Simon, F., S. Matheron, C. Tamalet, I. Loussert-Ajaka, S. Bartczak, J. M. Pépin, C. Dhiver, E. Gamba, C. Elbim, J. A. Gastaut, A. G. Saimot, and F. Brun-Vézinet.** 1993. Cellular and plasma viral load in patients infected with HIV-2. AIDS **7:**1411–1417.
60. **Spire, B., J. Sire, V. Zachar, F. Rey, F. Barré-Sinoussi, F. Galibert, A. Hampe, and J. C. Chermann.** 1989. Nucleotide sequence of HIV1-NDK a highly cytopathic strain of the human immunodeficiency virus HIV-1. Gene **81:**275–284.
61. **Srinivasan, A., D. York, D. Butler, Jr., R. Jannoun-Nasr, J. Getchell, J. McCormick, C. Y. Ou, G. Myers, T. Smith, E. Chen, G. Flaggs, P. Berman, G. Schochetman, and S. Kalyanaraman.** 1989. Molecular characterization of HIV-1 isolated from a serum collected in 1976: nucleotide sequence comparison to recent isolates and generation of hybrid HIV. AIDS Res. Hum. Retroviruses **5:**121–129.
62. **Swofford, D.** 1992. PAUP (phylogenetic analysis using parsimony) version 3.1. Center for Biodiversity, Illinois Natural History Survey, Sinauer Associates, Inc. Sunderland, Mass.
63. **Vanden Haesevelde, M., J.-L. Decourt, R. J. De Leys, B. Vanderborght, G. van der Groen, H. van Heuverswijn, and E. Saman.** 1994. Genomic cloning and complete sequence analysis of a highly divergent African human immunodeficiency virus isolate. J. Virol. **68:**1586–1596.
64. **Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon.** 1985. Nucleotide sequence of the AIDS virus, LAV. Cell **40:**9–17.
65. **Zekeng, L., L. Gurtler, E. Afane Ze, A. Sam-Abbenyi, G. Mbouni-Essomba, E. Mpoudi-Ngolle, M. Monny-Lobbe, J. B. Tapko, and L. Kaptue.** 1994. Prevalence of HIV-1 subtype O infection in Cameroon: preliminary results. AIDS **8:**1626–1628.
66. **Zwart, G., H. Langedijk, L. van der Hoek, J. J. de Jong, T. F. W. Wolfs, C. Ramautarsing, M. Bakker, A. de Ronde, and J. Goudsmit.** 1991. Immunodominance and antigenic variation of the principal neutralization domain of HIV-1. Virology **181:**481–489.
67. **Zwart, G., T. F. W. Wolfs, R. Bookelman, S. Hartman, M. Bakker, C. A. B. Boucher, C. Kuilen, and J. Goudsmit.** 1993. Greater diversity of the HIV-1 V3 neutralization domain in Tanzania compared with the Netherlands: serological and genetic analysis. AIDS **7:**467–474.