Virology **270,** 267–277 (2000)
doi:10.1006/viro.2000.0247, available online at http://www.idealibrary.com on IDEAL®

Independent Introduction of Transmissible F/D Recombinant HIV-1 from Africa into Belgium and the Netherlands

Eline Op de Coul,*[1] Audrey van der Schoot,† Jaap Goudsmit,† Remco van den Burg,† Wouter Janssens,‡ Leo Heyndrickx,‡ Guido van der Groen,‡ and Marion Cornelissen†

*Division of Public Health and Environment, Municipal Health Service, Nieuwe Achtergracht 100, 1018 WT Amsterdam, the Netherlands; †Department of Human Retrovirology, Academic Medical Center, Meibergdreef 45, 1105 BA Amsterdam, the Netherlands; and ‡Department of Microbiology, Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerp, Belgium

Received September 21, 1999; returned to author for revision January 10, 2000; accepted February 11, 2000

Most HIV-1 subtype F viruses described so far have been isolated from individuals originating in South America, Romania, or Central Africa. Previous studies have shown that subtype F viruses from these three areas can be distinguished by phylogenetic tree analysis of various parts of the HIV genome. Subtype F strains circulating in Central Africa and classified as subgroup F2 and F3 have relatively large nucleotide distances from strains of subgroup F1, which includes some African strains, along with strains from Romania and South America. Subtype F strains have now appeared in Europe. In this study, we analyzed the complete *gag* gene and a large fragment of the *pol* gene of seven strains of African origin that represent the three F subgroups. At least five of the seven strains appear to be intersubtype recombinants. Of four strains circulating in Belgium and the Netherlands, three were F/D mosaics and the fourth harboured a $G^{gag}/GH^{pol}/F3^{env}$ recombinant structure. Two of the three F/D mosaics showed identical breakpoints and were independently introduced in Belgium and the Netherlands. At least two of the mosaics were further transmitted. The remaining three strains of the seven we studied were isolated from individuals in Cameroon. Two included large or smaller F1 fragments in *gag* and *pol*. The third strain was subtype D along the entire *gag* and *pol* fragment. A parental African subtype F that showed no evidence for recombination was not found.

INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) strains of the major group M are subdivided into eight "pure" subtypes (A–D, F–H, J) and various circulating recombinant forms (CRFs) on the basis of phylogenetic analysis of partial and full-genome sequences (Korber *et al.,* 1998). Phylogenetic tree analyses reveal that the genetic subtypes of HIV-1 have nucleotide differences of 20–30% in their envelope (*env*) gene (Korber *et al.,* 1998). In most European countries, HIV-1 subtype B is the predominant subtype, but non-B subtypes have been introduced (Arnold *et al.,* 1995; Alaeus *et al.,* 1997; Lasky *et al.,* 1997).

HIV-1 subtype F is predominantly found in Central Africa, South America, and Romania (Dumitrescu *et al.,* 1994; Apetrei *et al.,* 1997). African countries where HIV-1 subtype F viruses have been detected are Cameroon (Delaporte *et al.,* 1996; Triques *et al.,* 1999), the Democratic Republic of Congo (DRC) (Mokili *et al.,* 1999; Triques *et al.,* 1999), Gabon, and the Central African Republic (CAR) (Triques *et al.,* 1999). Reported prevalences of *env* HIV-1 subtype F viruses in these African countries vary from 3% to approximately 10% (Janssens *et al.,* 1997; Takehisa *et al.,* 1998). In some European countries, sporadic cases of *env* HIV-1 subtype F viruses have been reported, mainly among individuals with an epidemiological link to Central Africa or South America. Recently, seven subtype F cases were described in Belgium, six of which were independent introductions between 1985 and 1994 and one was a mother-to-child transmission. At least two of these Belgium *env* subtype F strains presumably originated from West Central African countries (Heyndrickx *et al.,* 1998). HIV-1 *env* subtype F viruses have also been introduced in the Netherlands and strains of African origin have been isolated from individuals with an epidemiological link to the DRC (Wolfs *et al.,* 1992).

Authors of a recent paper on the genetic heterogeneity of HIV-1 subtype F strains from Central Africa (Triques *et al.,* 1999) apparently found certain strains difficult to classify. These strains are genetically closely related to HIV-1 subtype F, but form distinct clusters beyond the subtype F cluster first described, which included strains from Africa, Brazil, and Romania. To resolve the classification, the problematic strains were divided into three subgroups, F1, F2, and F3, based on comparison of their *env* and *gag* sequences. Group F1 includes strains from Romania, South America, and the DRC, whereas F2 and

[1] To whom reprint requests should be addressed at Division of Public Health and Environment, Municipal Health Service, Nieuwe Achtergracht 100, 1018 WT Amsterdam, the Netherlands. Fax: (31-20) 555-5533. E-mail: eodcoul@gggd.amsterdam.nl.

0042-6822/00 $35.00

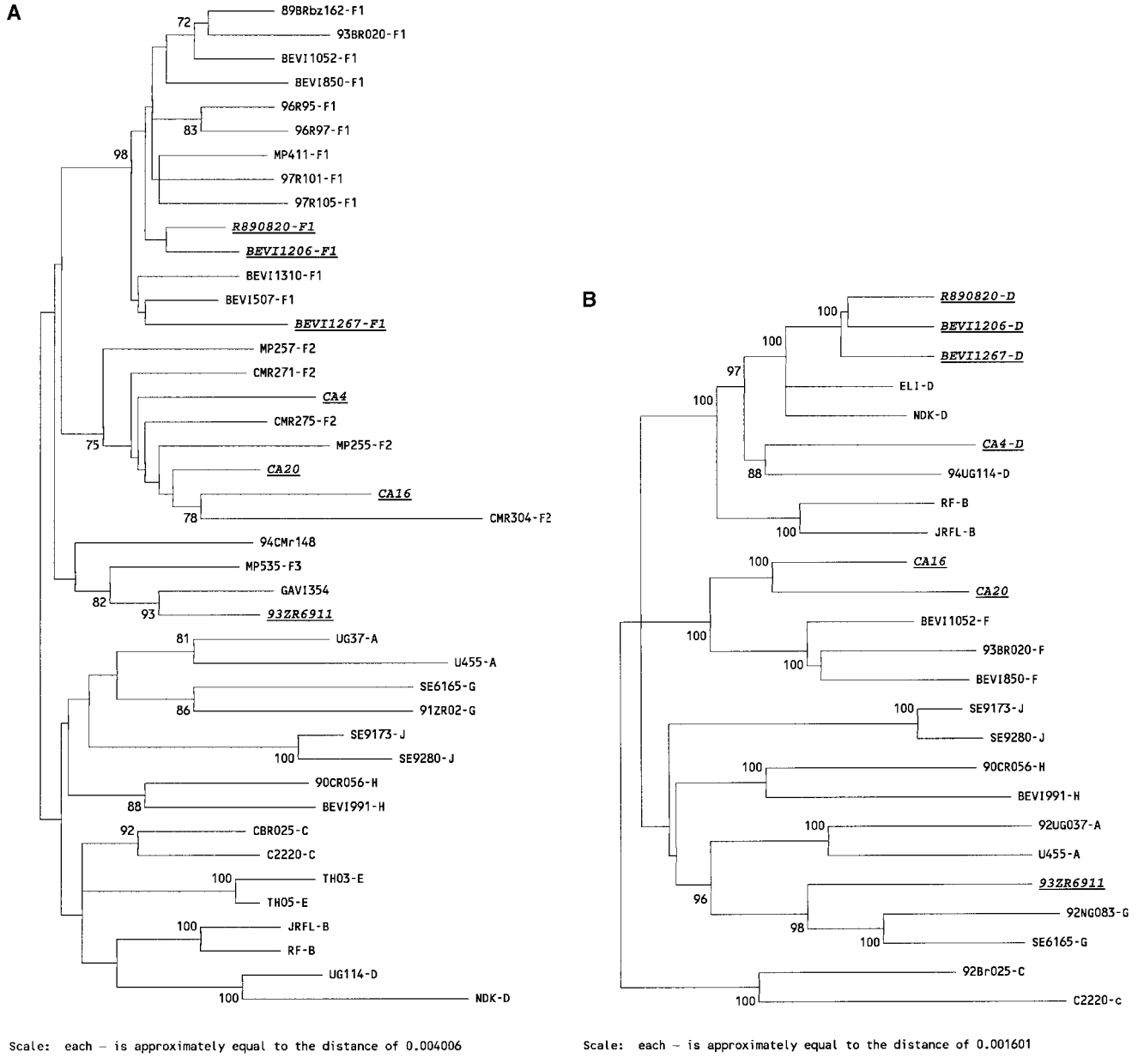

FIG. 1. Neighbour-joining phylogenetic trees of HIV-1 based on the V3 region of the *env* gene (276 nt, tree A), and the full genome of *gag* (1464 nt, tree B) and *pol* (1146 nt, tree C). The trees were constructed using a set of reference sequences from the Los Alamos GenBank: subtype A (U455, UG37); B (RF, JRFL); C (92Br025, C220); D (UG114, NDK); F (93BR020 from Brazil and BEVI850 from the DRC); G (SE6165), H (90CR056, BEVI991); and J (SE9173, SE9280). For these subtypes full-length sequences were available, but for E and I, only *env* sequences were available. Other references used are 89BRbz162 and BEVI1052 from Brazil, and 96R95, 96R97, 97R101, and 97R105 from Romania. The numbers in the figure represent the bootstrap value (100 replicates) for neighbour joining. Only bootstrap values higher than 70 are shown.

F3 are represented by African strains from Cameroon (F2, F3) and the DRC (F3). These subgroups have nucleotide differences of 23.5–27.8% in *env* and 6.8–9.7% in p24 *gag* and may represent new subtypes (Triques *et al.,* 1999). For most African strains classified as subtype F, sequence information for the *pol* gene is limited and for the *gag* gene mainly partial sequences were published. We therefore sequenced and phylogenetically analyzed the complete *gag* gene and a large *pol* fragment for several viruses that previously were identified as subtype F on the basis of their *env* gene (Wolfs *et al.,* 1992; Nkengasong *et al.,* 1994; Heyndrickx *et al.,* 1998; Op de Coul *et al.,* 1998). Four of these strains had been introduced into Belgium and the Netherlands, where at least two were transmitted further, one passing from a mother to her child and one passing to a sexual partner. We examined these strains to determine which viruses circulating in Belgium and the Netherlands are recombinant forms and which might represent the parental African subtype F. Furthermore, we examined three strains

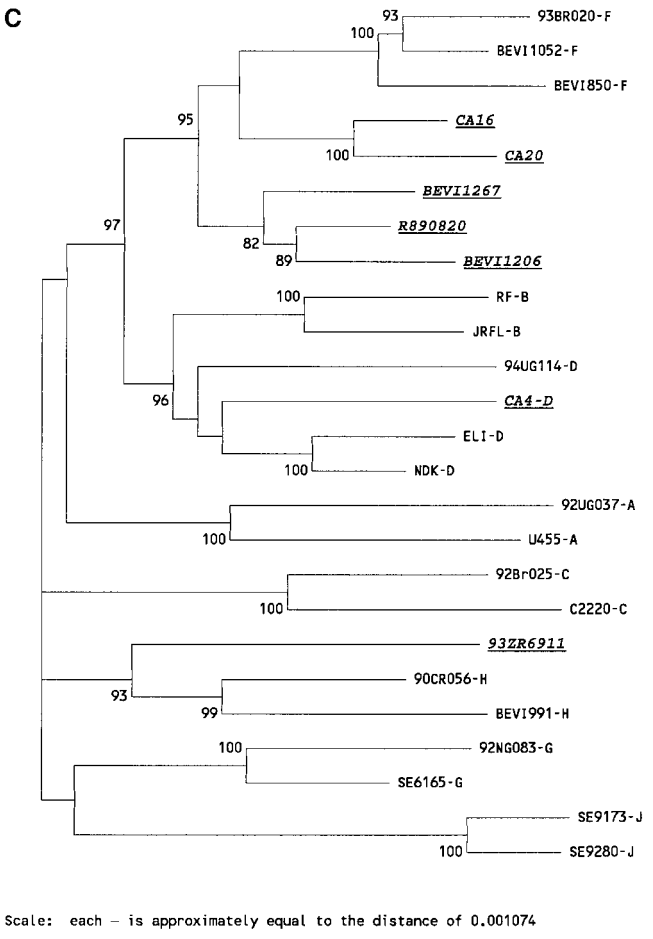

FIG. 1—*Continued*

from Cameroon that clustered as subtype F2 in order to understand the phylogenetic relations between strains from the subgroups F1, F2, and F3 of HIV-1.

RESULTS

Comparison with published African HIV-1 subtype F sequences

In order to phylogenetically characterize HIV-1 subtype F strains of African origin introduced into Belgium and the Netherlands and to understand the phylogenetic relations with published subtype F sequences from Africa, we sequenced the entire *gag* gene and a large portion of *pol* of seven HIV-1 strains that were previously classified as *env* subtype F (Nkengasong *et al.,* 1994; Janssens *et al.,* 1997; Op de Coul *et al.,* 1998). The samples were collected after we conducted a BLAST search (Altschul *et al.,* 1999) on an HIV-1 subtype F virus that was introduced in the Netherlands. In our previous study on HIV-1 subtype F viruses, in which we analyzed sequences from Romania, we observed that an HIV strain isolated from a Dutch man (R890820) who had a female partner from the DRC clustered in the phylogenetic tree within the group of Romanian subtype F sequences (Op de Coul *et al.,* 2000). By using the BLAST search this sequence was compared with the HIV sequence database to identify the most closely related African sequences. The R890820 *env* sequence showed a high homology of 94 and 92% with two African subtype F sequences (BEVI1206 and BEVI1267) identified in Belgium. In the *env* V3 phylogenetic tree in Fig. 1A it is shown that the three *env* sequences clearly belong to the genetic subtype F, but were placed beyond the Romanian and South American subclusters. HIV strain 93ZR6911, which was also introduced in the Netherlands, clustered in between the subtype F group that included reference sequences from Romania/South America and a group of Cameroon sequences (CA4, CA16, CA20, 94CMr304, 94CMr271, 94CMr275) which were also tentatively characterized as subtype F in the sequence database. Their distinct position from the Romanian/South American subtype F cluster was previously observed by other researchers and recently these sequences were classified as HIV-1 subtype F2 (Triques *et al.,* 1999). The BLAST search revealed that the 93ZR6911 strain showed highest homology with strain GAVI354 (94%) from Gabon (Delaporte *et al.,* 1996), also tentatively registered as subtype F in the sequence database. Furthermore, sequence 93ZR6911 showed homology with the African strains EQTB11, ZR36, MP535, and MP446, recently classified as subgroup F3 (Triques *et al.,* 1999). To this day, only one subtype F strain from sub-Saharan Africa has been described (BEVI850) that clusters consistently with subtype F references in phylogenetic trees based on the *env, gag,* and *pol* gene. This is due partly to the limited number of *pol* sequences available and partly to intersubtype recombination. To examine the possibility of intersubtype recombination among the subtype F strains introduced in Belgium and the Netherlands, we sequenced the complete *gag* gene (1464 nucleotides (nt)) and a large fragment of *pol* (1146 nt) of four strains, which clustered in *env* as F1 (R890820, BEVI1206, BEVI1267) and F3 (93ZR6911). Additionally, we analysed three sequences from Cameroon that clustered as subgroup F2 (CA4, CA16, and CA20).

The *gag* sequences of the three *env* subtype F1 strains—R890820, BEVI1206 and BEVI1267—were found to cluster together, as in the *env* tree, although this time with subtype D references (Fig. 1B). Furthermore, the *gag* sequences formed a distinct subcluster within the major subtype D cluster, with a bootstrap value of 100. One of the Cameroon sequences (CA4) also clustered within the D subtype but was highly divergent from the subtype D references included in the tree analysis. The *gag* sequence of 93ZR6911 also showed a discordant branching from the *env* tree. Although it clustered as subgroup F3 in *env,* it clustered as a highly divergent subtype G strain in *gag.* As in the *env* tree, the strains CA16 and CA20 did not clearly group within the clusters

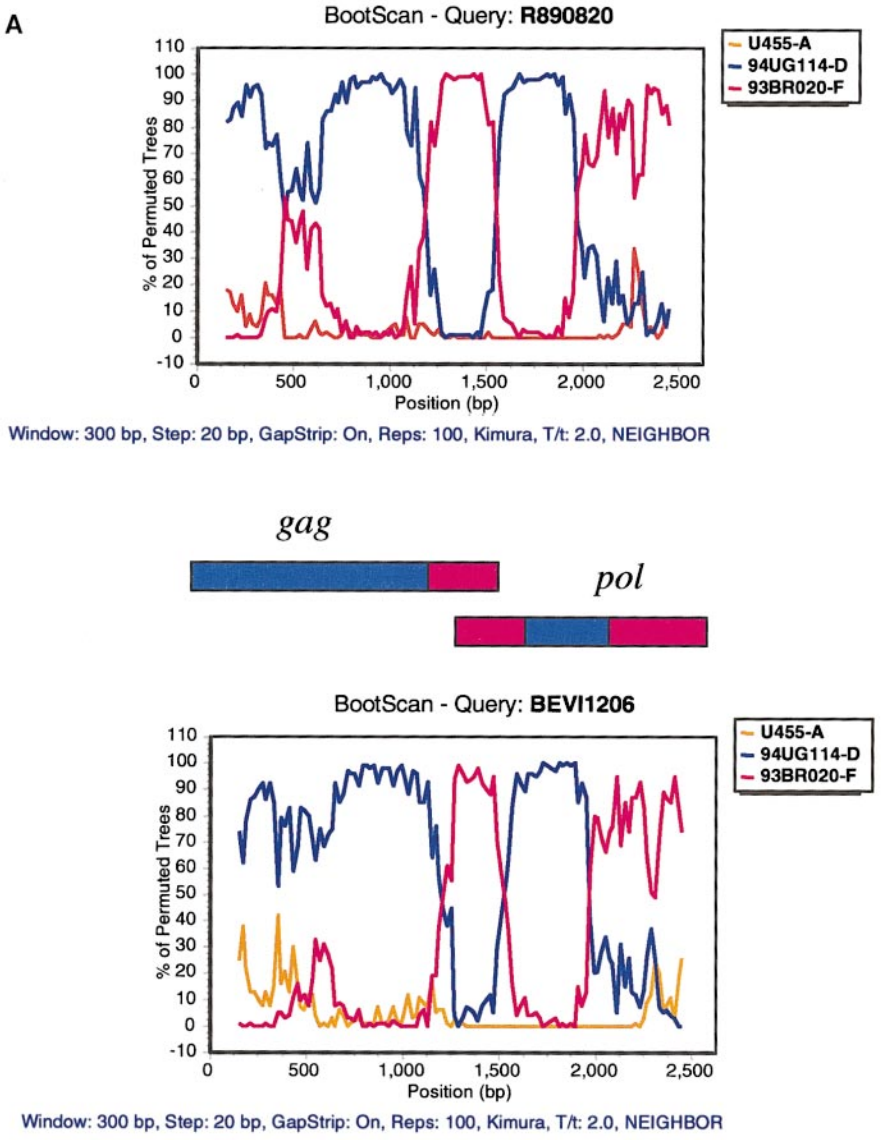

FIG. 2. Bootscan analysis of the complete *gag* and partial *env* and *pol* sequences of four African recombinant strains to determine the recombination breakpoints. The analysis is based on neighbour-joining tree analysis (Kimura two-parameter distances) with bootstrapping. The bootstrap values that support the clustering of the query sequences with the references are plotted. The chosen window size is 300 nt, moving with steps of 20 nt along the alignment. The recombination breakpoints are visualised in the bars above the graph. The unresolved parts of the genome are shown in white.

of *gag* reference sequences for the distinct subtypes A–J, although they appeared to have more affinity with the subtype F cluster than with other subtypes. We also constructed a *gag* tree based on P24 sequences of 624 nt in which we included the P24 sequences MP255 and MP257 obtained from the GenBank and classified by Triques *et al.* as subgroup F2 (data not shown). This tree showed that the four strains from Cameroon also formed one cluster in *gag,* but the bootstrap value of this cluster was low (49).

Although these Cameroon sequences were recently classified as a subgroup F2 of HIV-1 subtype F, their genetic relationship with the subtype F1 cluster is similar to the relationship previously observed for the subtypes B and D. In the phylogenetic tree based on the 1146 nt *pol* fragments (Fig. 1C), the three F1^{env}/D^{gag} recombinant strains R890820, BEVI1206, and BEVI1267 no longer belonged to subtype D, but were placed in a distinct subgroup with bootstrap value 82, close to the subtype F reference sequences, but with the sequences CA16 and CA20 in between. The *pol* fragment of strain 93ZR6911 clustered with bootstrap value 93 as a highly distinct variant of subtype H. The CA4 sequence remained subtype D in the tree based on *pol* sequences. So far, these results show that at least five out of seven African strains harbour intergenic recombinant structures.

To determine whether the sequences with interspersed positions or long branches represent intersub-

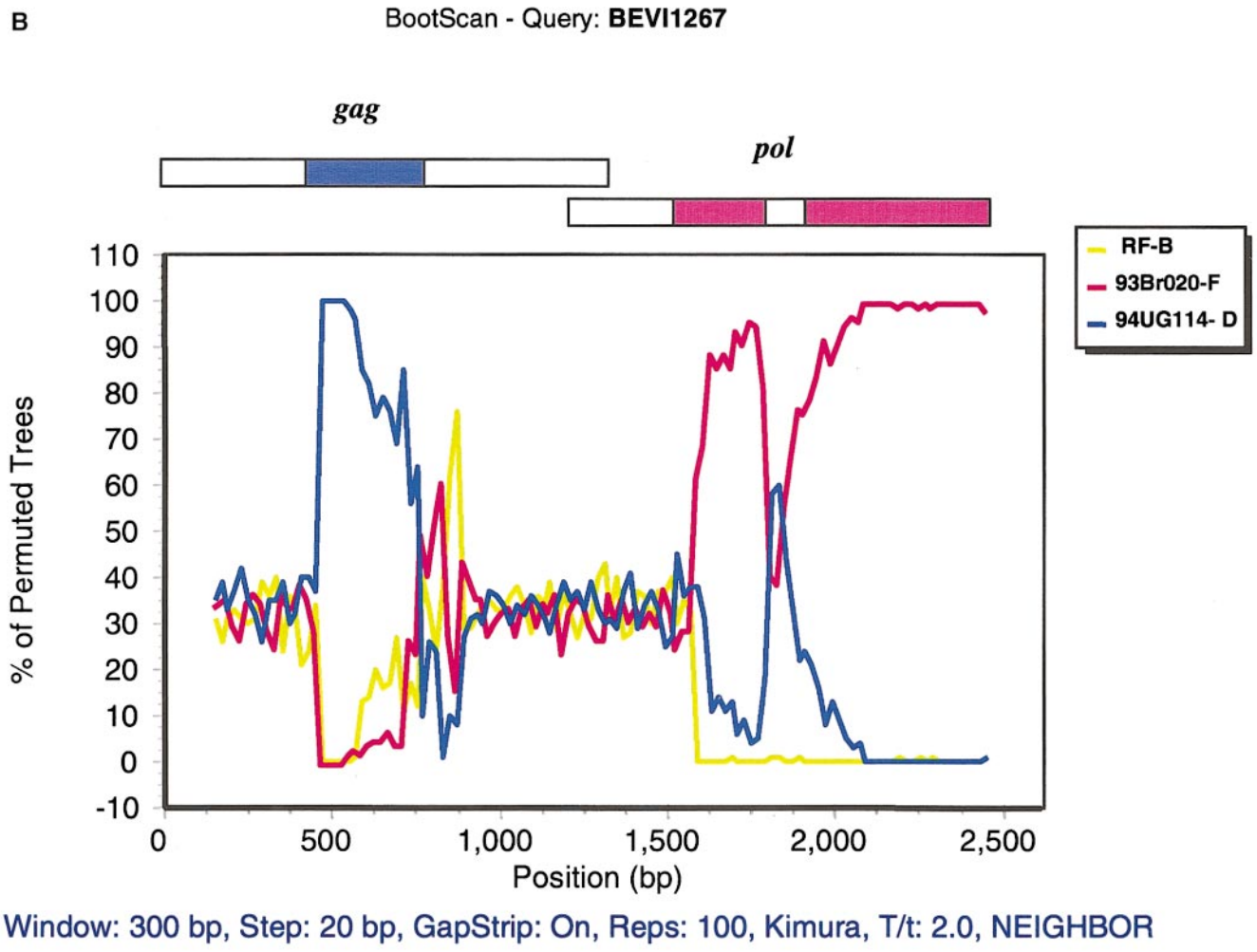

FIG. 2—*Continued*

type *gag* or *pol* intragenic recombinants, we constructed trees based on smaller overlapping segments of the *gag* and *pol* gene of approximately 400 nt each. We did this analysis for each putative recombinant sequence separately. This "short-fragment" tree analysis revealed that the first 400 nt of *gag* sequence 93ZR6911 clustered interspersed between the subtypes G and A, suggesting recombination between strains from the subtypes A and G. The first 400 nt of the *pol* sequence 93ZR6911 clustered interspersed between subtypes G and H, indicating that a recombination breakpoint is localised in this area. The second and third parts of 400 nt clustered consistently with the subtype H reference. The tree that included the first 400 nt of *pol* (protease and approximately 100 nt of RT) of the strains BEVI1206, BEVI1267, and R890820 showed positions interspersed between subtypes D and F. The second 400 nt fragment of the three strains clustered with subtype D, but the last fragment clustered again with subtype F. When the same analysis was conducted for strains CA16 and CA20, the sequences could not be assigned to any of the reference subtypes included in the *gag* (Fig. 1B) and *pol* tree (Fig. 1C), but consistently clustered with interspersed and similar positions that were relatively near but not within the subtype F cluster. These viruses may therefore belong to a "new" subtype, but the possibility of intersubtype recombination should be studied more extensively before such a conclusion is reached.

Localization of recombination breakpoints

The Recombination Identification Program (RIP) was used to confirm recombination and to determine all putative parents of the recombinant structures. The sequences were compared with reference sequences of each subtype: A through J. The *env* V3 sequences of the strains CA4, CA16, CA20, and 93ZR6911 and the *gag* sequences of CA16 and CA20, previously classified as F2 or F3, could not be assigned to any of the reference sequences included in the RIP analysis. There were no clear indications for recombination events in these regions. As expected, the *pol* sequences of R890820,

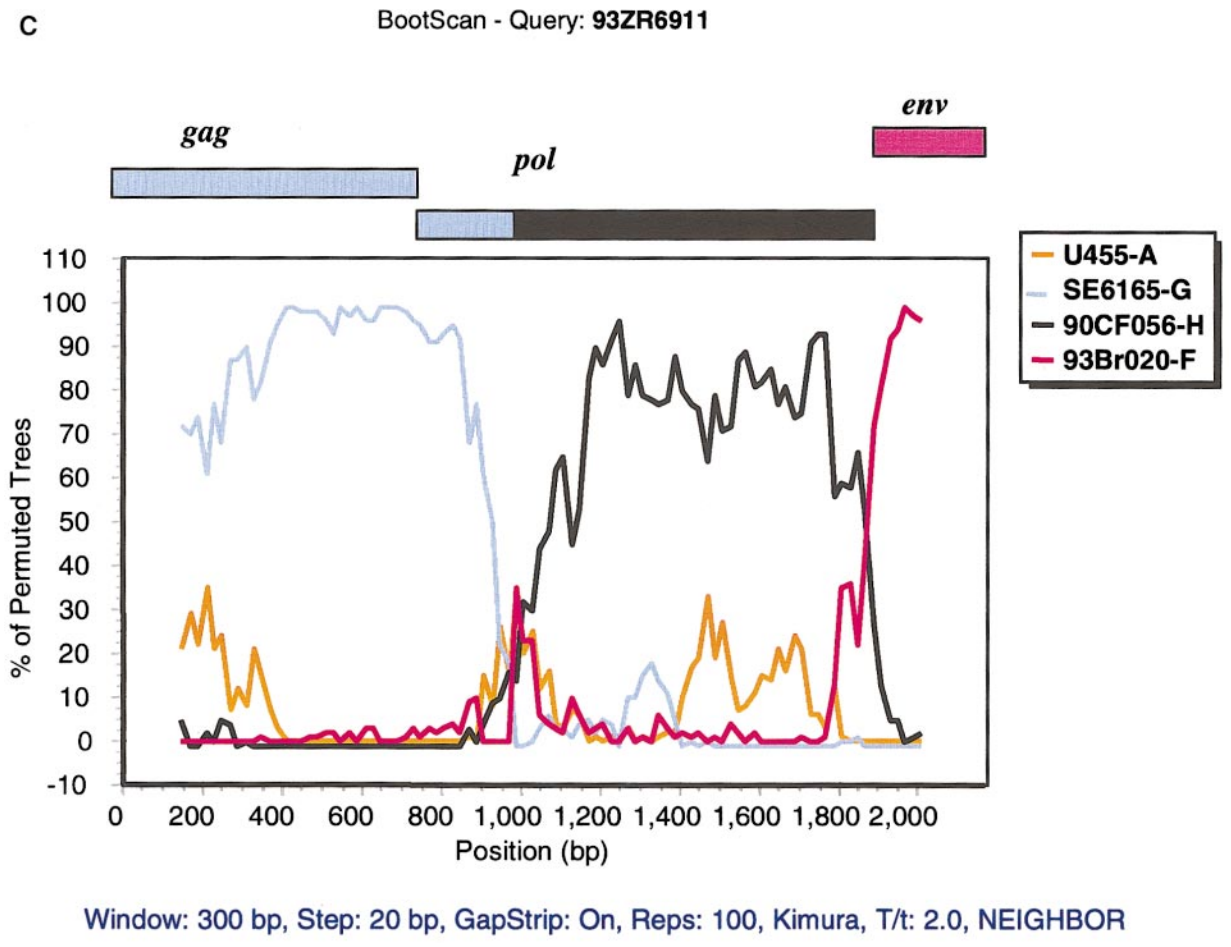

FIG. 2—*Continued*

BEVI1267, and BEVI1206 showed most similarity with the reference sequences for subtypes D and F. Strain 93ZR6911 (*gag* plus *pol* sequence) consisted of parts belonging to subtypes A, G, and H. The *pol* sequences of CA16 and CA20 contained fragments of subtype F, whereas the CA4 strain matched with the subtype D reference along the entire *gag* and *pol* sequence. To determine more precisely the breakpoints of the intersubtype recombinant sequences, we performed bootscan analysis (Ray, 1999) as described under Material and Methods. This analysis, in which a sliding window of 300 nt moved across the aligned sequences with steps of 20 nt, confirmed the results that we observed in the short-fragment tree analysis and the RIP analysis. The analysis was repeated with a window of 200 nt moving with steps of 10 nt with similar results (data not shown). The bootscan graph for the *gag* and *pol* sequence of strain R890820 (Fig. 2A) shows at least three recombination breakpoints, the first within the *gag* gene at about nucleotide position 1250, and second and third within the reverse-transcriptase (RT) gene in *pol.* Interestingly, the bootscan analysis revealed identical breakpoints in *pol* for BEVI1206 (Fig. 2A) and R890820, from Belgium and the Netherlands, respectively, and both represent individuals with an epidemiological link with the DRC. The $F1^{env}/D^{gag}/DF^{pol}$ recombinant sequences were independently introduced in the two countries and belonged to different individuals. Strain BEVI1267 (Fig. 2B) also contained parts of the subtypes D and F, but certain parts of *gag* and *pol* could not be assigned to any subtype of the HIV-1 group M. Strain 93ZR6911 (Fig. 2C) clustered with the subtype G reference in a 746 nt *gag* fragment, and in the first 250 nt of the *pol* gene. However, the rest of the *pol* sequence clustered with subtype H, whereas the *env* part showed most homology with the Brazilian subtype F reference. In contrast to the RIP analysis, the bootscan analysis did not provide evidence for a subtype A fragment in the *gag* gene. Figure 2D shows that sequence CA16 clustered with subtype F throughout the entire sequenced *gag* and *pol* fragment, although the bootscan value did not consistently reach the 70% level of significance. Partly this is due to the *gag-pol* transition area, which has a low sequence information density, and this causes a drop in the bootscan value at position 1100 in the graph. The *gag, pol,* and *env* regions of strain CA20 showed a complex and partly

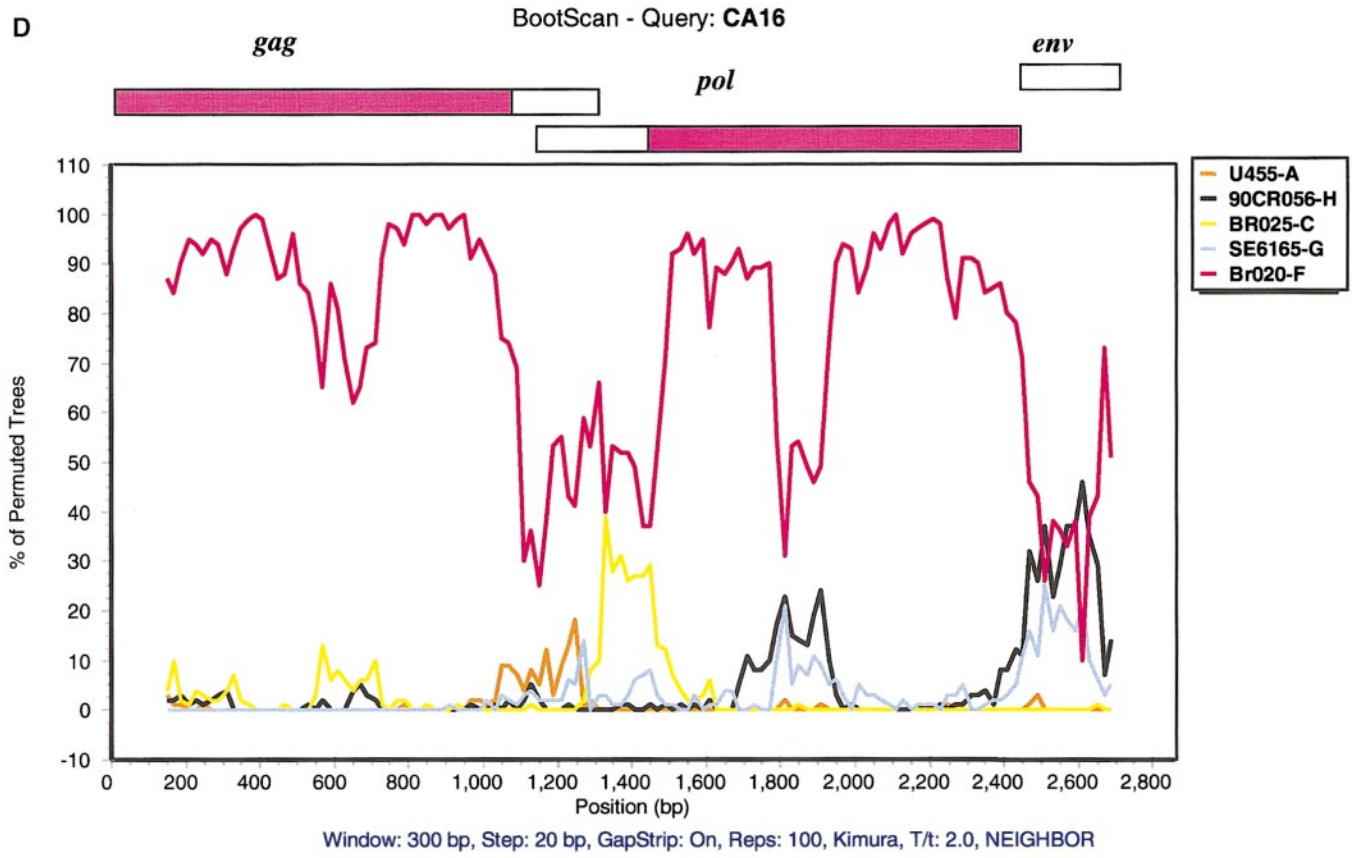

FIG. 2—*Continued*

unclear picture (not shown), but the first 400 nt of *gag* and the last 800 nt of *pol* clustered with subtype F. The 1400 nt in between these areas could not significantly be assigned to any subtype.

DISCUSSION

In our study of samples from seven individuals infected with the HIV-1 *env* subtype F virus, we sequenced the *gag* gene and an 1146 nt *pol* fragment in order to examine the HIV strains for intersubtype recombination. At least five of the seven HIV-1 strains, which represented subgroups F1, F2, and F3 (Triques *et al.,* 1999), appear to be intersubtype recombinants (R890820, BEVI1206, BEVI1267, CA4, 93ZR6911). Three strains (R890820, BEVI1206, and BEVI1267) are recombinants of subtypes F and D: subtype F in *env* and subtype D in *gag* with two differing structures in *pol.* Strains R890820 and BEVI1206, which were independently introduced in the Netherlands and Belgium, show identical breakpoints in the *pol* gene, halfway through the reverse-transcriptase gene. This crossover site was previously observed (Cornelissen *et al.,* 1996). To exclude the possibility that the strains with identical breakpoints have an epidemiological link we checked the birth dates of the female partner of R890820 and the Belgium woman BEVI1206 and they were different.

The strains CA4 and 93ZR6911 from Cameroon and the DRC also represent various intersubtype recombinants, with strain 93ZR6911 including at least three subtypes: G, H, and what appears to be F. The *env* part of this strain shows most homology with *env* regions of recently published strains (Triques *et al.,* 1999), classified as F3. However, based on genetic distances Triques *et al.* discuss the possibility that these regions could originate from a new subtype. Strains CA16 and CA20 from Cameroon appear to be intersubtype recombinants that most likely include segments of an unknown subtype, but they also include subtype F segments, which could explain its closer relationship with subtype F in the tree based on *gag* and *pol* fragments. The problematic sequences therefore do not fulfill all the criteria for a new subtype, since they do not consistently cluster in one group equidistant from the other subtypes in each genomic fragment.

At least five of seven HIV-1 subtype F viruses of African origin that were included in this study appear to be intersubtype recombinants. This could be due to sampling bias or to the relative rarity of HIV-1 subtype F in Africa and the presence of many other subtypes. The parental African subtype F virus was perhaps an unfit virus that could not compete with the many other HIV-1 subtypes circulating in Africa. The nonrecombinant sub-

TABLE 1

Epidemiological Information for Study Subjects

Sample	Gender	Origin of subject	Year of sampling	City of sampling	Risk exposure	HIV *env* subtype
R890820[a]	M	The Netherlands	1989	Amsterdam	Female partner from the DRC	F1
BEVI1206	F	Belgium	≤1994	Antwerp	Traveling in Africa	F1
BEVI1267[a]	F	Belgium	≤1994	Antwerp	Male partner from DRC	F1
93ZR6911	F	DRC	1993	Amsterdam	Heterosexual contact	F3
CA4	na[b]	Cameroon	1993	Yaoundé	na	F2
CA16	na	Cameroon	1993	Yaoundé	na	F2
CA20	na	Cameroon	1993	Yaoundé	na	F2

[a] Evidence of transmission of F/D recombinant strains.
[b] na, information is not obtained.

type E parent of the E/A recombinant strain is still unidentified, while the E/A recombinant strain is spreading fast in Asia. The high number of recombinants found among the relatively limited number of *env* subtype F viruses from Africa that have so far been sequenced suggests that this virus has increased its fitness, either by chance or as a survival mechanism, by recombining with other subtypes that are more readily and widely spread. In Brazil and in Argentina as well, recombinants and mosaics that include subtype F fragments have been reported (Sabino *et al.,* 1994; Marquina *et al.,* 1996). The spread of a B/C strain in China and the A/G-IbNg in sub-Saharan Africa (McCutchan *et al.,* 1999) strengthens the hypothesis that recombination does not have to be disadvantageous and may even promote the survival of the virus. Recombination is a genetic mechanism for the generation of multiple variants from single variants. It enables the virus to try a number of variants in its struggle with the host's immune system and eventually, the fittest variant will surpass the others. So far, there is no evidence that intersubtype recombination is harmful for the HIV-1 virus. In fact, it has been shown that recombinant viruses can readily spread further (McCutchan *et al.,* 1999). Two of the three F/D recombinants we studied were further transmitted: one strain was passed on from a mother to her child (BEVI1267) and another was sexually transmitted (R890820). The independent introduction of similar recombinant HIV strains (R890820 and BEVI1206) in two different countries could be coincidental, but could also point to the emergence of this particular recombinant strain. In that case, this F/D recombinant could become widely dispersed as time goes on.

In this study we observed intragenic crossover sites within the *gag* and *pol* gene, in addition to those previously described for the *env* (Neilson *et al.,* 1999), *gag,* and *pol* gene (Cornelissen *et al.,* 1996; Korber *et al.,* 1998) and the long terminal repeat (Blanckard *et al.,* 1999). The three F/D recombinant viruses exhibited not only a breakpoint between the *env* and the *gag* gene but also evidence for intragenic exchange within *pol* between subtype D and F strains. Intragenic recombination within the *pol* gene is not surprising, since this gene is highly conserved, and conserved parts of the genome are possibly more subject to recombination. The crossover site that we observed in *pol* was not localized at the protease-reverse transcriptase transition point but within the RT region. The *pol* fragments of the samples R890820 and BEVI1206 have similar breakpoints, suggesting a single shared crossover event in the past. Also possible, but perhaps less likely, is that the crossover between the subtypes D and F occurred at the same positions in distinct individuals who were coinfected with subtype D and F strains. The breakpoint in the *pol* gene of strain BEVI1267 has a different location than in the other strains, which implies that the exchange of sequence information between subtype D and F viruses has occurred at separate points during the course of the epidemic. The recombination event most likely occurred in Central Africa, where the prevalence of non-B subtypes is much higher than in Belgium or the Netherlands.

It is striking to note that recombination seems to occur more often between viruses that belong to the more closely related subtypes, such as subtype F and D or A, G, and H. This observation was not pursued by our study, but suggests the possibility that the strains with parts of more closely related subtypes are not true recombinants but consist of fragments with different evolutionary rates possibly caused by different immunological pressures. On the other hand, coinfections with distinct genotypic HIV strains have been described (Diaz *et al.,* 1995; Janini *et al.,* 1998). In one study even with evidence of viral recombination in an acute seroconverter (Zhu *et al.,* 1995).

Furthermore, the categorization of HIV-1 viruses as different subtypes or as subgroups of a subtype (as F1, F2, and F3) is basically a matter of arbitrary definition. It is possible that the F subgroups represent unidentified subtypes with a relatively small nucleotide distance.

In conclusion, this study demonstrates that three F/D recombinants from Africa have been introduced indepen-

dently in Belgium and the Netherlands. Secondary transmission of at least two of these recombinant viruses has occurred in both countries. The recent appearance of HIV-1 *env* subtype F viruses in Russia (Leitner *et al.,* 1996), the Philippines (Santiago *et al.,* 1998), Belgium (Fransen *et al.,* 1996), and the Netherlands (Op de Coul *et al.,* 1998) may suggest that subtype F is spreading. Whether its spread is the result of emerging recombinant strains is a hypothesis that cannot be confirmed without additional sequence analysis of other genes besides the *env.* Full-length sequence analysis is also needed to further characterise and possibly rename the strains now classified as F2 and F3. See Note added in proof.

MATERIALS AND METHODS

Study subjects

Seven samples seropositive for HIV-1 *env* subtype F were obtained for sequencing from the Department of Microbiology of the Institute of Tropical Medicine in Antwerp, Belgium, and the Academic Medical Center in Amsterdam, the Netherlands. Four of these samples had been taken from heterosexual individuals living in the Netherlands or Belgium. One of these (R890820) was obtained from a subject described in a previous study (Wolfs *et al.,* 1992), a Dutch man, who suffered from an acute HIV-1 infection in 1989 in the Netherlands. His HIV-seropositive female partner originated from the DRC. The three other samples came from persons whose strains were closely related to R890820, as determined by a BLAST search of the HIV sequence database. They included BEVI1267, taken from a Belgium woman who acquired the virus from a HIV-seropositive partner from the DRC and subsequently passed it to her child; sample BEVI1206, taken from a Belgium woman who most likely became infected in Africa; and sample 93ZR6911, obtained from a woman originally from the DRC who was living in Amsterdam (Op de Coul *et al.,* 1998). Additionally, we sequenced three *env* subtype F strains that were collected from heterosexual individuals in Cameroon (CA4, CA16, and CA20) (Nkengasong *et al.,* 1994; Heyndrickx *et al.,* 1998). The *env* sequences of the seven viruses described in the present study were previously published under Accession Numbers AJ228222, AJ228225, AF032169, ZR9306911, X80541, X80443, X80448. The 1464 nt *gag* and 1146 nt *pol* sequences were newly obtained for this study and submitted to the GenBank under Accession Numbers AF247517–AF247523. Table 1 summarizes the epidemiological information that was obtained for the seven study subjects.

DNA Sequencing and phylogenetic tree analysis

Viral RNA was extracted from 100 μl of serum (Amsterdam) or culture supernatant (Antwerp) (Boom *et al.,* 1990) and transcribed into complementary DNA (cDNA). The cDNA was used as templates for nested PCR amplification of sequences encoding almost the entire *gag* gene (p17, p24, p2, p7, p1, and p6), a total of 1464 nt, plus a large portion of *pol* (the complete protease and half of reverse transcriptase), a total of 1146 nt. The 276 nt V3 region of the *env* genes has been published previously (Wolfs *et al.,* 1992; Heyndrickx *et al.,* 1998). We conducted nested PCR amplifications for both the *gag* and the *pol* fragments. The *gag* gene was amplified in two parts of 729 nt and 735 nt, respectively. The *pol* gene was amplified in three partly overlapping fragments (A, 336 nt; B, 543 nt; and C, 487 nt) using primers and PCR conditions described in a previous paper (Cornelissen *et al.,* 1996). The amplified double-stranded DNA products were directly sequenced in two directions on the ABI 377 automatic sequencer (Applied Biosystems, Foster City, CA) using the dye terminator primer set. The sequences were aligned by using Clustal X (Thompson *et al.,* 1994) and subsequently revised manually. The nucleotide alignments were included in phylogenetic trees generated by the neighbour-joining method (MEGA program) (Kumar *et al.,* 1993) using Kimura two-parameter distances. The trees included reference sequences according to Los Alamos (Korber *et al.,* 1998) representing different HIV-1 subtypes for which there is no evidence for intersubtype recombination: subtype A (U455, UG37), B (RF, JRFL), C (92BR025, C2220), D (94UG114, NDK), F (92BR020), G (SE6165), H (90CF056, BEVI991), and J (SE9173, SE9280) (Salminen *et al.,* 1996; Carr *et al.,* 1998; Gao *et al.,* 1998; Korber *et al.,* 1998; Laukkanen *et al.,* 1999). Subtype E and I sequences were included only in the phylogenetic tree analysis of the *env* gene, since the *gag* and *pol* parts of these viruses show mosaic structures with subtypes A and A/G, respectively (Carr *et al.,* 1998). Bootstrap analysis was conducted to test the phylogenetic clusters for statistical significance.

Examination for intersubtype recombination

All sequences were analysed with the Recombination Identification Program to examine the possibility of intersubtype recombination in the *env, gag,* and *pol* fragments (Siepel *et al.,* 1995). The sequences were compared with reference sequences for the subtypes A to J. When recombination was demonstrated, a more extensive analysis was conducted by phylogenetic tree analysis of short genomic fragments. Subsequently, the sequences were scanned by an advanced technique of bootscanning, as implemented in the SIMPLOT program for Microsoft Windows, in which a panel of reference sequences (sliding window) moves across the query sequence (Ray, 1999). The bootscan analyses were conducted with various window sizes and window steps. In each step, a phylogenetic tree with 100 replicates was constructed by NEIGHBOR, from the PHYLIP package, using Kimura two-parameter distances of DNADIST

(Felsenstein, 1997). Sequence gaps were excluded from the analysis. The bootstrap value of each query sequence and its corresponding reference sequence were plotted. Reference sequences that showed no relationship with the recombinant strains were removed, and the analysis was repeated to simplify the graphic figure and to determine the breakpoints more precisely. This second bootscan analysis was done only when removing the sequences did not significantly influence the shape of the graph and the bootscan values. The breakpoints were localized as the midpoints of the transition from one subtype to another. The bootscan analysis was conducted separately for the *gag* and *pol* genes (data not shown), since their information density varies because of differences in mutation variability. The *env* part could not be analysed separately, given the shortness of the fragment. The bootscan analysis was therefore conducted for the whole sequence, including the *gag, pol,* and *env* fragments. Since intragenic differences in nucleotide variability could disturb the result of the bootscan analysis, we adjusted the window size from 200 to 300 nt and the corresponding number of window steps (from 20 to 10 nt) to see if the picture would change. The bootscan analysis was also conducted in an effort to characterise genomic fragments that could not be assigned to any of the reference sequences in the phylogenetic trees. Fragments shorter than 200 nt and having a bootstrap value lower than 70% were considered to be not significant and therefore unresolved.

ACKNOWLEDGMENTS

We thank Margreet Bakker for providing data for this study, Roel Coutinho and Volodya Lukashov for their critical reading of the manuscript, and Lucy Phillips for her editorial work. This study was financially supported by the "Ziekenfondsraad" in the Netherlands as part of the Stimulation Program on AIDS Research on the Dutch Program Committee for AIDS Research (Grant VWS 96-02/96.1006) and by the "Fonds voor Wetenschappelijk Onderzoek" (Grants 3-3301-96 and G-0134-97).

Note added in proof. The F3 virus is recently renamed as subtype K in *AIDS Res. Hum. Retroviruses* (2000), **16**(2), 139–151.

REFERENCES

Alaeus, A., Leitner, T., Lidman, K., and Albert, J. (1997). Most HIV-1 genetic subtypes have entered Sweden. *AIDS* **11,** 199–202.

Altschul, S. T., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402; BLAST Version 2.0.9, 1999 (http://www.ncbi.nlm.nih-.gov).

Apetrei, C., Loussert-Ajaka, I., Collin, G., Letourneur, F., Duca, M., Saragosti, S., Simon, F., and Brun-Vézinet, F. (1997). HIV type 1 subtype F sequences in Romanian children and adults. *AIDS Res. Hum. Retroviruses* **13,** 363–365.

Arnold, C., Barlow, K. L., Parry, J. V., and Clewly, J. P. (1995). At least five HIV-1 sequence subtypes (A,B,C,D,A/E) occur in England. *AIDS Res. Hum. Retroviruses* **3,** 427–429.

Blackard, J. T., Renjifo, B. R., Mwakagile, D., Montano, M. A., Fawzi, W. W., and Essex, M. (1999). Transmission of human immunodeficiency type 1 viruses with intersubtype recombinant long terminal repeat sequences. *Virology* **254,** 220–225.

Boom, R., Sol, C. J. A., Salimans, M. M. M., Jansen, C. L., Wertheim-van-Dillen, P. M. E., and Noordaa van der, J. (1990). Rapid and simple method for purification of nucleic acids. *J. Clin. Microbiol.* **28,** 495–503.

Carr, J. K., Salminen, M. O., Albert, J., Sanders-Buell, E., Gotte, D., Birx, D. L., and McCutchan, F. E. (1998). Full genome sequences of human immunodeficiency virus type 1 subtypes G and A/G intersubtype recombinants. *Virology* **247,** 22–31.

Cornelissen, M., Burg van den, R., Zorgdrager, F., Lukashov, V. V., and Goudsmit, J. (1997). *Pol* gene diversity of five human immunodeficiency virus type 1 subtypes: Evidence for naturally occurring mutations that contribute to drug resistance, limited recombination patterns, and common ancestry for subtypes B and D. *J. Virol.* **71,** 6348–6358.

Cornelissen, M., Kampinga, G., Zorgdrager, F., Goudsmit, J., and the UNAIDS Network for HIV isolation and characterization (1996). Human immunodeficiency virus type 1 subtypes defined by env show high frequency of recombinant gag genes. *J. Virol.* **70,** 8209–8212.

Delaporte, E., Janssens, W., Peeters, M., Buve, A., Dibanga, G., Perret, J. L., Ditsambou, V., Mba, J. R., Courbot, M. C., Georges, A., Bourgeois, A., Samb, B., Henzel, D., Heyndrickx, L., Fransen, K., van der Groen, G., and Larouze, B. (1996). Epidemiological molecular characteristics of HIV infection in Gabon, 1986-1994. *AIDS* **10,** 903–910.

Diaz, R., Sabino, E. C., Mayer, A., Mosley, J. W., Busch, M. P., and the transfusion safety study group (1995). Dual immunodeficiency virus type 1 infection and recombination in a dually exposed transfusion recipient. *J. Virol.* **69,** 3273–3281.

Dumitrescu, D., Kalish, M. L., Klicks, S. C., Bandea, C. I., and Levy, J. A. (1994). Characterization of human immunodeficiency virus type 1 isolates from children in Romania: Identification of a new envelope subtype. *J. Infect. Dis.* **169,** 281–288.

Felsenstein, J. PHYLIP (Phylogeny Inference Package, version Phylip95.exe). Department of Genetics, University of Washington, Seattle, WA (http://evolution.genetics.washington.edu/phylip/general.html).

Fransen, K., Buvé, A., Nkengasong, J. N., Laga, M., and van der Groen, G. (1996). Longstanding presence in Belgians of multiple non-B subtypes. *Lancet* **347,** 1403.

Gao, F., Robertson, D. L., Carruthers, C. D., Morrison, G., Jian, B., Chen, Y., Barré-Sinoussi, F., Girard, M., Srinivassen, A., Abimiku, G., Shaw, M., Sharp, P. M., and Hahn, B. H. (1998). A comprehensive panel of near-full-length clones and reference sequences for non-B isolates of human immunodeficiency virus type 1. *J. Virol.* **72,** 5680–5698.

Heyndrickx, L., Janssens, W., Coppens, S., Vereecken, K., Willems, B., Fransen, K., Colebunders, R., Vandenbruaene, M., and van der Groen, G. (1998). HIV type 1 C2V3 env diversity among Belgian individuals. *AIDS Res. Hum. Retroviruses* **14,** 1291–1296.

Janini, L. M., Tanuri, A., Schechter, M., Peralta, J. M., Vicente, A. C., Dela Torre, N., Pieniazek, N. J., Luo, C. C., Ramos, A., Soriano, V., Schochetman, G., Rayfield, M. A., and Pieniazek, D. (1998). Horizontal and vertical transmission of human immunodeficiency virus type 1 dual infections caused by viruses of subtypes B and C. *J. Infect. Dis.* **177,** 227–231.

Janssens, W., Buve, A., and Nkengasong, J. N. (1997). The puzzle of HIV-1 subtypes in Africa. *AIDS* **11,** 705–712.

Korber, B., Kuiken, C., Foley, B., *et al.* Human retroviruses and AIDS 1998. A compilation and analysis of nucleic acid and amino acid sequences. Los Alamos Natl. Lab., Los Alamos, NM.

Kumar, S., Tamura, K., and Wei, M. (1993). MEGA: Molecular evolutionary genetics analysis, version 1.0. Institute of Molecular Evolutionary Genetics, the Pennsylvania State Univ. Park, PA.

Lasky, M., Perret, J. L., Peeters, M., *et al.* (1997). Presence of multiple non-B subtypes and divergent subtype B strains of HIV-1 in individuals infected after overseas deployment. *AIDS* **11,** 43–51.

Laukkanen, T., Albert, J., Liitsola, K., Green, S. D., Carr, J. K., Leitner, T., McCutchan, F. E., and Salminen, M. O. (1999). Virtually full-length sequences of HIV type 1 subtype J reference strains. *AIDS Res. Hum. Retroviruses* **15,** 293–297.

Leitner, T., Korovina, G., Marquina, S., Smolskaya, T., and Albert, J. (1996). Molecular epidemiology and MT-2 cell tropism of Russian HIV type 1 variant. *AIDS Res. Hum. Retroviruses* **12,** 1595–1603.

Marquina, S., Leitner, T., Rabinovich, R. D., Benetucci, J., Libonatti, O., and Albert, J. (1996). Coexistence of Subtypes B, F, and a B/F *env* Recombinant of HIV Type 1 in Buenos Aires, Argentina. *AIDS Res. Hum. Retroviruses* **12,** 1651–1654.

McCutchan, F. E., Carr, J. K., Bajani, M., Sanders-Buell, E., Harry, T. O., Stoeckli, T. C., Robbins, K. E., Gashau, W., Nasidi, A., Janssens, W., and Kalish, M. L. (1999). Subtype G and multiple forms of A/G intersubtype recombinant human immunodeficiency virus type 1 in Nigeria. *Virology* **254,** 226–234.

Mokili, J. L. K., Wade, C. M., Burns, S. M., Cutting, W. A. M., Bopoli, Green, S. D. R., Peutherer, J. F., and Simmonds, P. (1999). Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural Democratic Republic of Congo. *AIDS Res. Hum. Retroviruses* **7,** 655–664.

Neilson, J. R., John, G. C., Carr, J. K., Lewis, P., Kreiss, J. K., Jackson, S., Nduati, R. W., Mbori-Ngacha, D., Panteleeff, D. D., Bodrug, S., Giachetti, C., Bott, M. A., Richardson, B. A., Bwayo, J., Ndinya-Achola, J., and Overbaugh, J. (1999). Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya. *J. Virol.* **73,** 4393–4403.

Nkengasong, J. N., Janssens, W., Heyndrickx, L., *et al.* (1994). Genotypic subtypes of HIV-1 in Cameroon. *AIDS* **10,** 1405–1412.

Op de Coul, E. L. M., Lukashov, V. V., Doornum van, G. J. J., Goudsmit, J., and Coutinho, R. (1998). Multiple HIV-1 subtypes present amongst heterosexuals in Amsterdam 1988-1996: No evidence for spread of non-B subtypes. *AIDS* **12,** 1253–1255.

Op de Coul, E., van den Burg, R., Asjö, B., Goudsmit, J., Cupsa, A., Pascu, R., Usein, U., and Cornelissen, M. (2000). Genetic evidence of multiple transmissions of HIV type 1 subtype F within Romania from adult blood donors to children. *AIDS Res. Hum. Retroviruses* **16,** 327–336.

Ramos, A., Tanuri, A., Schechter, M., Rayfield, M. A., Hu, D. J., Cabral, M. C., Bandea, C. I., Baggs, J., and Pieniazek, D. (1999). Dual and recombinant infections: An integral part of the HIV-1 epidemic in Brazil. *Emerg. Infect. Dis.* **1,** 65–74.

Ray, S. C. (1999). Simplot for windows (version 2.5) Baltimore, MD (http://www.med.jhu.edu/deptmed/sray/download/simplot_doc.html).

Sabino, E. C., Shpaer, E. G., Morgado, M. G., Korber, B. T., Diaz, R. S., Bongertz, V., Cavalcante, S., Galvao-Castro, B., Mullins, J. I., and Mayer, A. (1994). Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals in Brazil. *J. Virol.* **68,** 6340–6346.

Salminen, M. O., Johansson, B., Sonnerborg, A., Ayehunie, S., Gotte, D., Leinikki, P., Burke, D. S., and McCutchan, F. E. (1996). Full-length sequence of an ethiopian human immunodeficiency virus type 1 (HIV-1) isolate of genetic subtype C. *AIDS Res. Hum. Retroviruses* **12,** 1329–1339.

Santiago, M. L., Santiago, E. G., Hafalla, J. C., Manalo, M. A., Orantia, L., Cajimat, M. N., Martin, C., Cuaresma, C., Dominguez, C. E., Borromeo, M. E., De Groot, A. S., Flanigan, T. P., Carpenter, C. C., Mayer, K. H., and Ramirez, B. L. (1998). Molecular epidemiology of HIV-1 infection in the Philippines, 1985 to 1997: Transmission of subtypes B and E and potential emergence of subtypes C and F. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **18,** 260–269.

Siepel, A. C., and Korber, B. T. (1995). Recombination Identification Program (RIP, version 1.3), Human retroviruses and AIDS 1997: A compilation and analysis nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, NM (http://hiv-web.lanl.gov/RIP).

Takehisa, J., Zekeng, L., Ido, E., Mboudjeka, I., Moriyama, H., Miura, T., Yamashita, M., Gurtler, L. G., Hayami, M., and Kaptue, L. (1998). Various types of HIV mixed infections in Cameroon. *Virology* **245,** 1–10.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (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; Clustal X (http://evolution.bmc.uu.se/~thomas/mol_linux/clustalx).

Triques, K., Bourgeois, A., Saragosti, S., Vidal, N., Mpoudi-Ngole, E., Nzilambi, N., Apetrei, C., Ekwalanga, M., Delaporte, E., and Peeters, M. (1999). High diversity of HIV-1 Subtype F strains in Central Africa. *Virology* **259,** 99–109.

Wolfs, T. F., Zwart, G., Bakker, M., and Goudsmit, J. (1992). HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. *J. Virol.* **189,** 103–110.

Zhu, T., Wang, N., Carr, A., Wolinsky, S., and Ho, D. D. (1995). Evidence for coinfection by multiple strains of human immunodeficiency virus type 1 subtype B in an acute seroconvertor. *J. Virol.* **69,** 1324–1327.