Journal of Medical Virology 81:391–399 (2009)

Genetic Variability of Human Immunodeficiency Virus-1 in Bahia State, Northeast, Brazil: High Diversity of HIV Genotypes

Joana Paixão Monteiro,[1] Luiz Carlos Junior Alcantara,[1,2] Tulio de Oliveira,[3] Antonio Marcos Oliveira,[1] Marco Antônio Gomes Melo,[1,2] Carlos Brites,[4] and Bernardo Galvão-Castro[1,2]*

[1]*Laboratório Avançado de Saúde Pública, CPqGM/FIOCRUZ, Salvador, Bahia, Brazil*
[2]*Escola Bahiana de Medicina e Saúde Pública, EBMSP, Salvador, Bahia, Brazil*
[3]*South African National Bioinformatics Institute, Cape Town, South Africa*
[4]*Universidade Federal da Bahia, Salvador, Bahia, Brazil*

The HIV-1 genetic variability in Bahia state, Brazil, was investigated. DNA samples from 229 and 213 HIV-1-infected individuals were analyzed using the heteroduplex mobility assay (HMA) in *gag* and *env* fragments, respectively. One hundred seventy-five samples were characterized in both genes. Thirty-two subtype F and BF recombinant viruses were sequenced and analyzed by phylogenetic methods. The combination of HMA and sequencing results showed that seven different HIV-1 genotypes comprised this sample: 147 (84%) B/B, 4 (2.3%) F/F, 3 (1.7%) B/F, 1 (0.6%) F/B, 1 (0.6%) F/D, 1 (0.6%) BF/F, and 18 (10.3%) BF/B. A significant divergence was observed between these two techniques results (84.4%). This is explained by the low accuracy of the HMA for detecting recombinant viruses. These recombinants were unrelated to CRF12, while two sequences were related to CRF28 and CRF29. Nineteen BF mosaics shared the same *gag* breakpoint. In conclusion, the use of HMA may be inappropriate in regions where different subtypes are co-circulating. Subtype B is the most common genotype, however, an increased prevalence (13.1%) of different BF variants and a potentially new CRF suggest that recombination is occurring frequently in Bahia. These viruses were associated with women infected heterosexually. Finally, this study identified the presence of an F/D recombinant HIV-1 in Brazil. ***J. Med. Virol. 81:391–399, 2009.***

KEY WORDS: HIV-1; epidemiology; phylogeny; Brazil; Bahia

INTRODUCTION

The *human immunodeficiency virus* (HIV) is characterized by high genetic variability and is classified into different types, groups, subtypes, and circulating recombinant forms (CRFs) [Mccutchan et al., 2000; Casado et al., 2005; Thomson et al., 2005; De Sa Filho et al., 2006], each presenting distinct distribution in the world. Moreover, HIV-1 genotypes show different dispersion among the different risk groups and transmission routes [Weniger et al., 1994; Kunanusont et al., 1995; Burke and McCutchan, 1996]. Nevertheless, the association between this genetic variability and viral features as transmission, infectivity, and pathogenesis is not elucidated to date.

Subtype B is the predominant HIV-1 genotype in Brazil. However, an increasing number of non-B and recombinants infections are being identified [Morgado et al., 2002; Eyer-Silva et al., 2007; Monteiro et al., 2007]. Within the Brazilian vast territory, different subtype distribution patterns have been reported among its five geographic regions. The southern and southeastern regions have been largely characterized in regard to HIV-1 molecular epidemiology. However, little information is available about other regions of the country. Particularly, in Bahia state, few previous studies have focused on the prevalence of HIV-1 subtypes [Couto-Fernandez et al., 1999; Pedroso et al., 2007] which considered only short fragments of the HIV genome, thus, probably under estimating the presence of mosaic virus.

Grant sponsor: Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB); Grant sponsor: Brazilian Ministry of Health STD/AIDS Program; Grant sponsor: CNPq.

*Correspondence to: Bernardo Galvão-Castro, Laboratório Avançado de Saúde Pública, Centro de Pesquisa Gonçalo Muniz, Fundação Oswaldo Cruz, Rua Waldemar Falcão 121, Candeal, Salvador, Bahia 40296-610, Brazil.
E-mail: bgalvao@bahia.fiocruz.br

Accepted 4 November 2008
DOI 10.1002/jmv.21414
Published online in Wiley InterScience (www.interscience.wiley.com)

Bahia is the fourth most populous Brazilian state and presents the highest number of notified cases of HIV/AIDS in the Northeast region. In this part of the country, economic and education privation, sexual tourism, and prostitution can be related to an increase in the incidence of AIDS and the introduction of new genetic forms is expected. Therefore, the identification of these variants is of great importance for HIV surveillance as well as for vaccine development and efficacy. This study was conducted in order to access the HIV genetic variability in Salvador, the capital of Bahia, to determine the prevalence of subtypes and to verify the association with groups and transmission routes. Also, the current molecular epidemiological profile of the HIV population in this area was compared with previous data from 10 years ago. Thus, besides its potential implications for therapeutic strategies at a local level, this study contributes to the understanding of the evolution of HIV.

This study has received official and ethical approval of the Centro de Pesquisa Gonçalo Moniz, Fiocruz, Bahia.

METHODS

Study Population

As part of a sectional study in 2002, blood samples from 261 HIV-1-seropositive subjects, unrelated epidemiologically and followed regularly at Professor Edgard Santos University Hospital, Salvador, Bahia, were obtained after informed consent. The samples were sent to the Advanced Laboratory of Public Health (LASP)/CPqGM/FIOCRUZ for further processing. Peripheral blood mononuclear cells (PBMCs) were separated by centrifugation on Ficoll-hypaque gradient. DNA was extracted from PBMCs using the Qiagen extraction kit (Qiagen, Valencia, CA). Clinical and epidemiological information were obtained from medical records.

Genotyping

Semi-nested PCR was conducted for individual amplification of *gag* and *env* fragments. The heteroduplex mobility assay (HMA) was based on *gag* and *env* fragments amplified by PCR. H1G777/H1P202 and H1Gag1584/G17 were used as outer and inner primer sets for *gag*, while for *env*, ED5/ED12 and ED31/ED33 were used as outer and inner primers, respectively. The PCR conditions and the HMA subtyping were carried out as described previously [Delwart et al., 1993; Heyndrickx et al., 2000]. Samples classified as subtype F in HMA*gag* and/or HMA*env* had further PCR amplification. The gag p17 (~1,150 bp) nested PCR protocol was carried out as described previously [Guimarães et al., 2002]. The same HMA primer set (~550 bp) or, alternatively, ES7/ES8 (~700 bp) were used for *env* amplification.

Sequencing and Analysis

PCR products were purified in Qiagen columns (Qiagen) and sequenced in an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) using the Big Dye Terminator kit (Applied Biosystems). Sequences were assembled using SeqMan software (DNAStar, Madison, WI) and reported to GenBank under the accession numbers EU770698–EU770752.

An alignment with a subtype reference set from Los Alamos database (http://hiv-web.lanl.gov) was created using CLUSTAL X software [Thompson et al., 1997] and edited manually in GENEDOC [Nicholas et al., 1997]. The *gag* and *env* fragments corresponded to positions 952–1958 and 6845–7324 relative to HXB2 genome, respectively. All the sequences were checked for contamination by BLAST search against the HIV-1 sequences database and among themselves. Neighbor-joining (NJ) and maximum likelihood (ML) trees were reconstructed in the PAUP* 4.0b10 software [Swofford, 1999], using the appropriate nucleotide substitution model implemented in the Modeltest software [Posada and Krandall, 1998]. The reliability of each cluster was determined using 1,000 bootstrap replicates. For the ML tree, a heuristic search was undertaken with a subtree-pruning-regrafting branch swapping algorithm using the NJ tree as the starting material, including its optimized parameters. The likelihood ratio test (RT) method was used to calculate statistical support for the branches: $P < 0.001$ (highly significant **) and $P < 0.005$ (significant *). Bootstrap and ML supports were added to NJ tree that was drawn with the TreeView 1.4 software [Page, 1996].

The Bootscanning method implemented in the SIMPLOT software [Salminen et al., 1995] was used to investigate the intra-gene recombination within *gag* fragments, and the GENEDOC software was used to determine the specific crossover point based on the HXB2. Sequences were fragmented at this breakpoint and the segments were submitted to phylogenetic analysis with Los Alamos reference data sets. Intra-subtype relationships were investigated through individual phylogenetic analysis of subtypes B, F, and D *gag* and *env* sequences (including new sequences, Brazilian sequences characterized previously, and sequences from other countries) as described above. Samples had their *gag* and *env* sequences concatenated (1,500 bp) and submitted to phylogenetic analysis using the method described above. The between and within-group genetic distances were measured using the Kimura 2-α-parameter model with a distance matrix implemented in the MEGA 4.0 package [Tamura et al., 2007]. The standard error computation was obtained by bootstrap analysis (1,000 replicates).

RESULTS

In this sectional study, HMA analyses were conducted in order to identify the HIV-1 subtype of 155 men (59.4%) and 106 (40.6%) women. Out of these, 228 and 208 samples were individually analyzed in *gag* and *env* fragments, respectively. Based on *gag* analysis, 202 (88.6%) samples were classified as subtype B, 25 (11.0%) as subtype F, and 1 (0.4%) as subtype D, while in *env*

region, 204 (98.1%) samples were classified as B, 3 (1.4%) as F, and 1 (0.5%) as D. Subtype F was at least 3 times more frequent in *gag* than in *env* gene. One hundred seventy-five samples were characterized in both genomic regions; pure subtype B was assigned for 143 (81.7%) samples and pure subtype F for 1 (0.6%). Three recombinant forms were identified: 6 (3.4%) B*gag*/F*env*, 24 (13.7%) F*gag*/B*env*, and 1 (0.6%) D*gag*/B*env*. Concerning the association between HIV-1 subtypes and gender, 64% (16) of the samples classified as subtype F in *gag* were obtained from women, while 63.4% (128) of the B samples were isolated from men. In *env*, all three F samples represented men and 40.7% (83) of the B samples represented women. Regardless of the genomic segment analyzed, subtype B was the most prevalent among both genders.

In attempt to characterize the subtype F viruses circulating in Brazil, *gag* (952–1958 relative to HXB2) and *env* (6845–7324 relative to HXB2) fragments of the 32 samples classified as subtype F in the *gag* and/or the *env* HMA were sequenced and submitted to phylogenetic analyses (Fig. 1). Some samples are not shown in the trees, either because they had shorter *gag* sequences, were amplified with different *env* primers or have not been amplified in *env* region. In the NJ analysis of *env*, 18 sequences clustered as subtype B, 3 as subtype F, and 1 as subtype D. In *gag* analysis, 7 sequences grouped with subtype B references, 6 sequences grouped within the subtype F reference cluster, and 18 sequences formed a separated but closely related to the F group (bootstrap = 70). Therefore, these sequences were examined further using the Bootscanning recombination tool. All these sequences and one other smaller sequence (109) presented a recombination pattern between subtypes B and F in the *gag* fragment (Fig. 2a), with the crossover site located between positions 1370 and 1410 (relative to HXB2). In order to determine if these BF recombinants were phylogenetic related, the exact recombination point was sought through the inspection of subtype signature nucleotides in comparison to the Los Alamos reference set with the GENEDOC software. All the BF recombinant sequences shared the same breakpoint, located at p24 (position 1398 in HXB2). NJ and ML analyses of the individual fragments using the appropriate evolutionary model confirmed the Bootscanning recombination results (Fig. 2b).

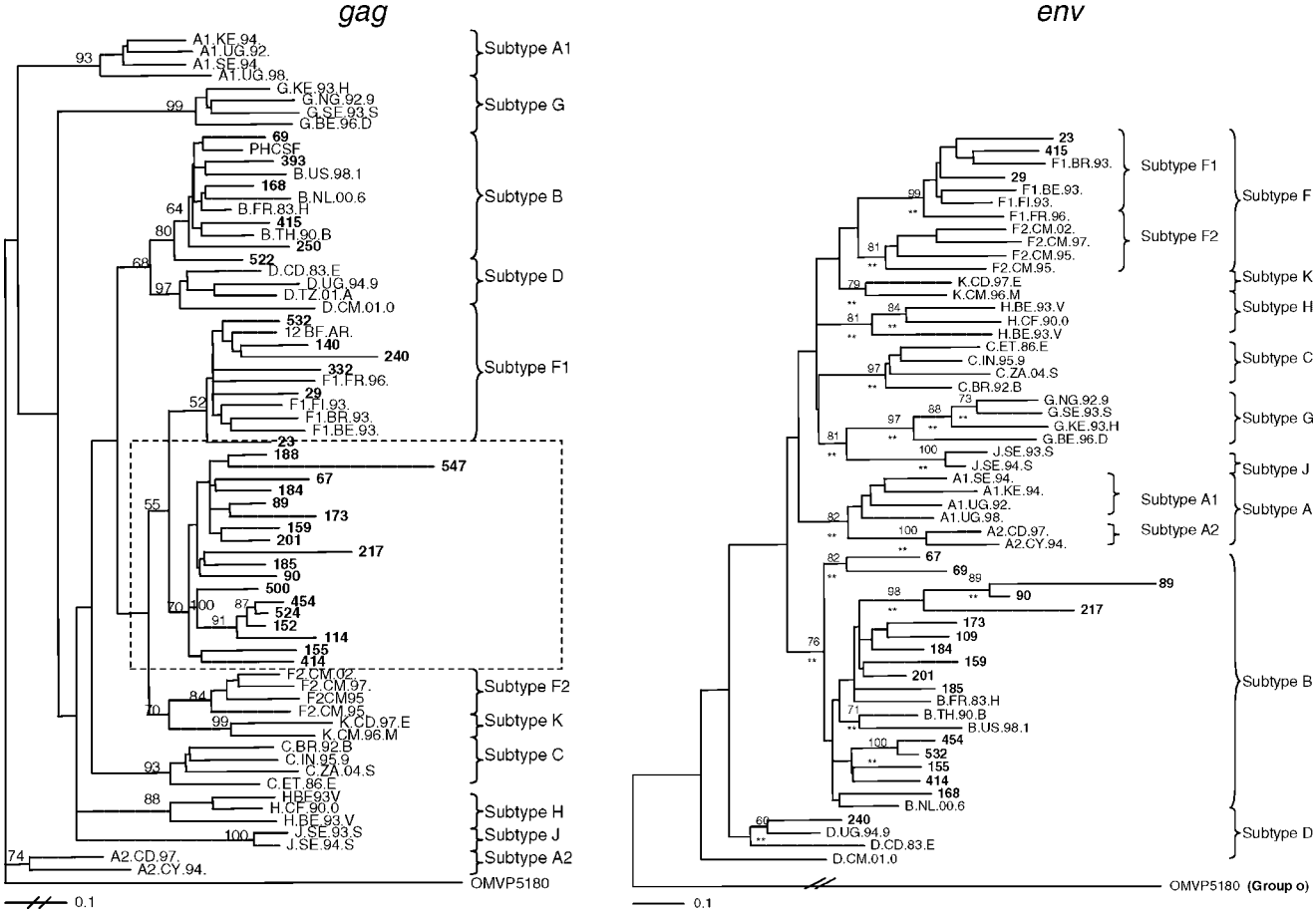

Fig. 1. *gag* and *env* phylogenetic relationships among HIV-1 samples from Bahia and group M references sequences from Los Alamos database. Group O sequences were used as outgroup. Trees were constructed based on the NJ method using the K81uf + I + G (*gag*) and the GTR + I + G (*env*) substitution models, respectively. Bootstrap values for 1,000 replicates are indicated. Branches highly supported by the ML method are indicated as ** ($P < 0.001$). HIV-1 sequences generated in the present study are indicated in bold. Eighteen *gag* sequences formed a unique monophyletic group outside the subtype F1 cluster shown within the box.

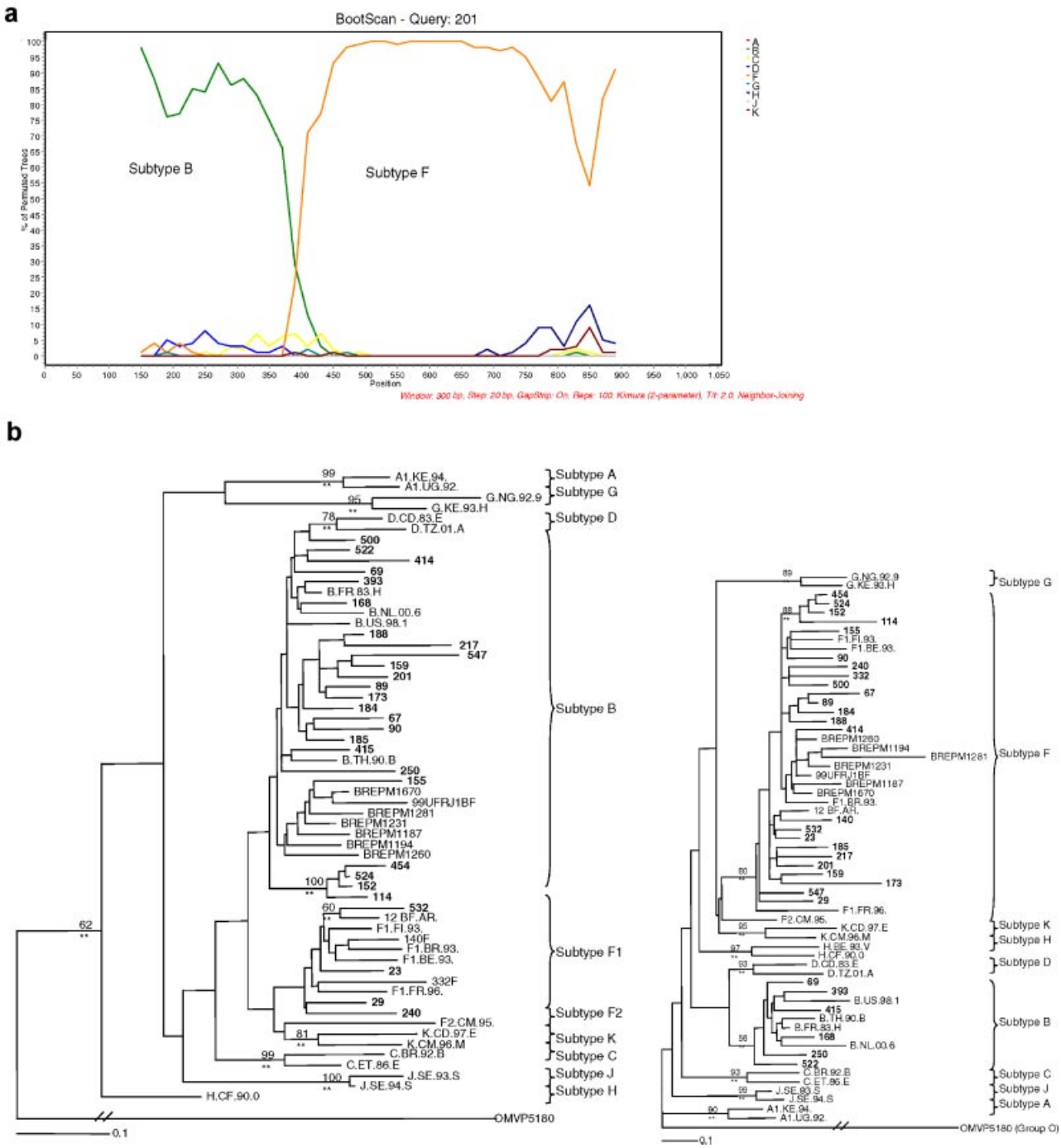

Fig. 2. BF recombinants viruses among Brazilian HIV-1 samples. **a**: Bootscan analysis showing the recombinant pattern between subtypes B and F inside the *gag* gene. **b**: Sequences were fragmented at the breakpoint and the segments (HXB2 nucleotides 952–1397 and 1398–1961) were submitted to NJ analysis with Los Alamos reference set using the F81uf + I + G and the TRN + I + G substitution models, respectively. Bootstrap values for 1,000 replicates are indicated. Branches highly supported by the ML method are indicated as ** ($P < 0.001$). HIV-1 sequences generated in the present study are indicated in bold.

The HMA and phylogenetic results (Table I) were compared. The following discordance was found: based on *gag*, two samples classified as subtype F (393 and 522) and one as subtype D (168) with HMA were indeed subtype B, while three samples classified as subtype B grouped within the subtype F (29 and 140) and BF (524) clusters. Based on *env*, two samples classified as subtype B with HMA were actually one subtype F (23) and one subtype D (240) in the phylogenetic tree and one classified as subtype F was inside the subtype B group (69).

Thus, considering the combined HMA and sequencing analyses (Table I), the 175 samples were classified in

TABLE I. Clinical and Epidemiological Characteristics and Genetic Subtypes of HIV-1-Infected Individuals From Bahia

	Subtype[a]		Clinical and epidemiological data						
Sample	*gag*	*env*	Gender	Age	Diagnosis (year)	Transmission route	Viral load (copies/ml)	CD4 (cells/ml)	AIDS symptoms
23	F	F	M	38	NA	NA	NA	NA	Y
29	F	F	M	38	NA	NA	NA	NA	N
67	BF	B	F	NA	NA	Ht	NA	NA	NA
89	BF	B	M	38	1996	Ht	1,600	176	N
90	BF	B	M	39	1997	Hm	Ldl	631	N
109	BF	B″	F	NA	NA	Ht	NA	NA	NA
114	BF	B	F	NA	1999	Ht	4,700	85	N
140	F	F	F	29	1997	Ht	Ldl	543	N
152	BF	B	F	28	1997	Blood	Ldl	781	N
155	BF	B	M	40	1997	Ht	2,300	694	N
159	BF	B	F	36	1996	Ht	8,700	210	Y
173	BF	B	M	25	1999	Ht	450,000	233	N
184	BF	B	F	23	1996	Ht	250,000	65	Y
185	BF	B	F	NA	NA	Ht	NA	NA	NA
188	BF	B	F	NA	NA	Ht	NA	NA	NA
201	BF	B	M	35	NA	NA	66,000	156	NA
217	BF	B	F	41	2000	IDU	<40	157	N
239	B	F	F	30	1999	Blood	<80	223	N
240	F	D	F	38	1998	Ht	110	246	NA
250	B	F	F	NA	NA	IDU	61,000	407	N
332	F	F	M	NA	2001	NA	1,700	393	N
414	BF	B	M	39	2002	Ht	10,000	303	N
415	B	F	F	NA	NA	Ht	NA	NA	NA
454	BF	B	F	NA	NA	Ht	NA	NA	NA
500	BF	B	F	39	NA	Ht	Ldl	612	NA
524	BF	F	F	35	NA	Ht	Ldl	549	N
532	F	B″	F	35	1996	Ht	Ldl	550	N
547	BF	B	F	NA	NA	Ht	NA	NA	NA

B″, Brazilian subtype B; NA, not available; M, male; F, female; Ht, heterosexual; Hm, homosexual; IDU, intravenous drug user; Ldl, lower than detectable levels; Y, yes; N, no.
[a]*gag* and *env* subtypes were defined based on phylogenetic analyses. For the samples that did not have enough material for the sequencing, results presented here are based on HMA.

seven different HIV-1 genotypes based on *gag*/*env* genomic regions: 147 (84%) B/B, 4 (2.3%) F/F, 3 (1.7%) B/F, 1 (0.6%) F/B, 1 (0.6%) F/D, 1 (0.6%) BF/F, and 18 (10.3%) BF/B. Table I summarizes the subtypes and clinical-epidemiological data of the 28 individuals infected with subtype F or F recombinant viruses. Overall, 67.9% were women, 79.2% reported heterosexual behavior, 8.3% were intravenous drug users, and 2 women (8.3%) were infected through blood transfusion. One man reported homosexual relation as the transmission route. Mean age was 35 years, median CD4 count was 303 cells/ml (interquartile range = 349) and median viral load was 1,600 copies/ml (interquartile range = 10,000). Most of the patients (83.3%) did not present symptoms typical of AIDS status.

To further characterize the intra-subtype relationships of the viruses circulating in this population and to analyze the relationships between new and previously described BF recombinants, phylogenetic trees were constructed using Los Alamos reference sequences of subtypes B, F, and D and other pure-subtype and recombinant sequences from South America (Fig. 3). In the *gag* full-segment analysis (Fig. 3a), like in the subtyping tree (Fig. 1), the 18 recombinant sequences clustered as a monophyletic group outside the subtype F1 cluster together with other BF recombinants described previously including sequences of two CRFs identified in southeastern Brazil [De Sa Filho et al., 2005, 2006]. However, two different monophyletic subclusters can be observed within this BF cluster: one formed by 16 of the recombinant sequences identified in this study and other BF recombinant sequences from Brazil and Argentina and the other formed by two of the new recombinant samples, CRF28 and CRF29 sequences. Moreover, inside the former cluster, four sequences (454, 524, 152, and 114) formed a separate cluster supported by a bootstrap of 90. The same topology for this group was observed in the trees based on individual subtype fragments of *gag* (Fig. 2b). None of the recombinants were closely related to the CRF12, although one F sequence (532) was in both inter (Fig. 1) and intra-subtype (Fig. 3) analyses of *gag*. The intra-subtype phylogenetic analysis of the envelope sequences (Fig. 3b) showed that subtype B sequences clustered inside two major groups. The first one was represented by all the CRF28 and CRF29 sequences in addition to five of the new sequences, all of which presenting the BF recombination pattern in *gag*. The second group was formed by nine samples obtained in this study, all the other sequences from Bahia [Couto-Fernandez et al., 1999] and other B sequences from other parts of Brazil and from other countries. All of

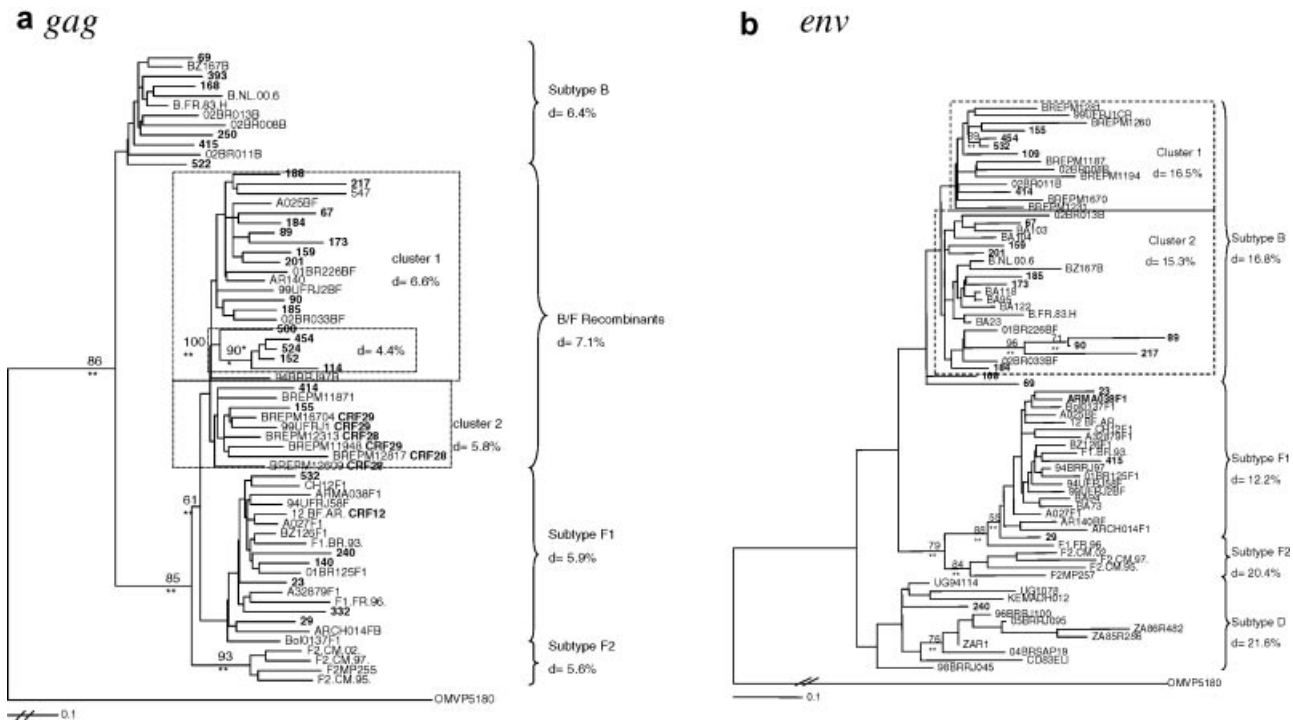

Fig. 3. **a**: *gag* and (**b**) *env* phylogenetic relationships among HIV BF recombinants from Bahia, compared to other subtypes B, F, and BF recombinant sequences from South America and around the world. Trees were constructed based on the NJ method using the K81uf + I + G (*gag*) and the GTR + I + G (*env*) substitution models, respectively. Bootstrap values for 1,000 replicates are indicated. Branches highly supported by the ML method are indicated as ** ($P < 0.001$). HIV-1 sequences generated in the present study are indicated in bold.

these nine subtype B sequences were BF recombinants in the *gag* fragment. Concerning the intra-subtype F relationships, the three new envelope sequences were dispersed among the vast reference group; however, none of them was closely related to the two other subtype F sequences from Bahia identified previously neither to the subtype F sequence of CRF12. The *env* subtype D sequence identified here clustered inside the subtype D group, however, this sequence was not related to any of the other sequences of subtype D identified so far in Brazil as well as to any other African sequence used as reference. The genetic distances for the subtypes and recombinant clusters are shown in the trees.

Diverse unique BF recombinant genomes as well as two CRFs have been described in other geographic regions of Brazil [De Sa Filho et al., 2005, 2006; Sanabani et al., 2006]. Thus, the genetic structures of such viruses were compared to the recombination patterns from Bahia (Fig. 4). In the *gag* and *env* concatenated tree, 19 new viruses had both genome fragments available. Out of these, all BF_{gag}/B_{env} and F_{gag}/B_{env} recombinants involved in this analysis formed a unique monophyletic group (bootstrap = 73). Again, two different recombinant clusters, with significant bootstrap values (84 and 90, respectively), were observed: one formed by CRF28 and CRF29 sequences in addition to two of the samples and the other composed by nine new recombinants and one recombinant from Brazil identified previously. The other BF recombinant pattern virus ($415 = B_{gag}/F_{env}$) clustered closer to subtype B sequences, however, outside the group. The FD recombinant sequence did not cluster with any pure-subtype or recombinant sequence.

DISCUSSION

The identification of HIV-1 genotypes represents an important tool for molecular epidemiology studies, helping the evaluation of changes in the HIV epidemic and also contributing to the understanding of the determinants of virus evolution. The HIV genetic diversity and distribution are well documented for the southern and southeastern regions of Brazil; however, few reports have focused on the remaining geographic regions. The aim of this study was to evaluate the HIV-1 subtype distribution in Salvador, the capital of Bahia state and the third most populous city in the country. With almost 3 million people living in the city, 13,750 are estimated to be infected with HIV in Salvador [Dourado et al., 2007]. The HMA testing was used to screen the subtypes in 261 HIV-positive samples since this technique has been described as a useful tool to determine HIV subtypes [Pereira et al., 2004; Rios et al., 2005; Cabral et al., 2006]. However, when the sequences of 32 subtypes F and BF recombinants were analyzed, a significant proportion of divergence was found between these methods (84.4%). These contrasting results can be attributed to the limited ability of the HMA to

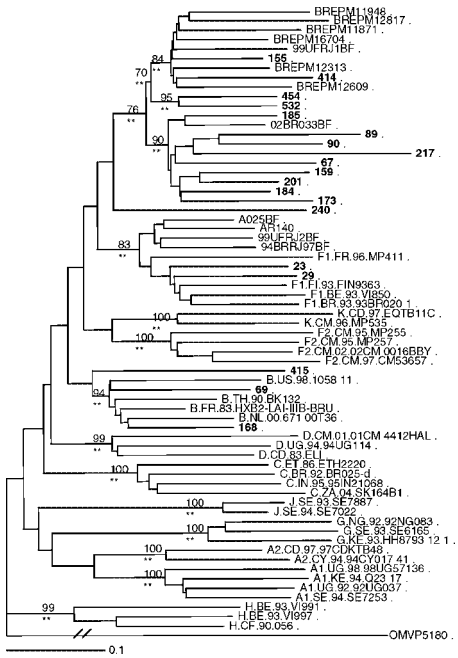

Fig. 4. NJ tree based on the concatenated *gag* and *env* regions showing the phylogenetic relationships of the Brazilian BF recombinants viruses relative to other subtypes and recombinant forms. Trees were constructed using the TVM + I + G substitution model. A group O sequence was used as outgroup. Branches highly supported by the ML method are indicated as ** ($P < 0.001$). HIV-1 sequences generated in the present study are indicated in bold.

distinguish BF recombinants as 19 intra-gag BF recombinants were identified as either subtype F or B isolates. Therefore, the use of the HMA may be inappropriate to identify HIV genotypes in regions where different subtypes co-circulate and where the presence of recombinant viruses is expected. In this context, Buonaguro et al. [2005] had proposed a modified version of HMA in which instead of being used for subtype classification, the technique is used to screen HIV-1 isolates divergent from the local predominant subtype by testing samples only against the reference standard representing that predominant subtype. Eighty-four percent of the studied samples were pure B, 2.3% were pure F, 13.1% were represented by different patterns of subtypes B and F recombinants, and 0.6% (one sample) was an F/D recombinant. Subtype C was not represented in this population. The differences in subtypes B and BF recombinants prevalences compared to the previous data (92% and 5.5%, respectively) [Couto-Fernandez et al., 1999] can be attributed to two different factors: (1) an increase in the circulation of BF recombinants and (2) a previous sub-estimation of the real prevalence of BF recombinants since only one genetic region was analyzed in that report. On the other hand, both present and previous studies reported a similar lower rate of pure subtype F (1.6%) compared to BF recombinants in Bahia. The 2.3% of subtype F found here are also lower than the prevalence found for the southeastern region, which varied from 3.7% to 9% [Guimarães et al., 2002; Couto-Fernandez et al., 2005; De Sa Filho et al., 2005; Barreto et al., 2006; Cabral et al., 2006]. Similarly, different reports based on samples from Argentina and Brazil have shown that most of the samples characterized previously as subtype F in one genomic region were indeed BF recombinants in the complete genome, suggesting that the occurrence of pure subtype F is rare [Carr et al., 2001; Guimarães et al., 2002; Thomson et al., 2002; Quarleri et al., 2004; Barreto et al., 2006]. If these mosaic viruses present evolutionary advantages over pure subtype F strains is a subject for future investigation. In this regard, out of the 23 recombinant viruses identified in this work and among five different patterns, 19 had subtype B sequence in *env*, which could be an indication of selective pressure to the maintenance of the subtype B envelope. Supporting this theory, one previous study [De Sa Filho et al., 2005] showed that among six different BF recombination patterns, all of them were subtype B in the envelope and another one [Monteiro et al., 2007] reported that from three different BC mosaic patterns, recombination always occurred with subtype B in *env*. On the contrary, these strains could be representing a founder effect event, deriving from a common recombinant ancestral that went through different evolution process and subsequent recombination.

The epidemiological data of the individuals enrolled in this analysis revealed that while subtype B epidemic follows the national tendency [Boletim Epidemiológico AIDS, 2007] for the general population with a male/female ratio of 2:1, the subtype F and F recombinants showed the opposite (sex ratio = 1:1.8) with 68% (19 against 9) of the infected patients being women (Table I). Also, almost 80% were infected heterosexually. These observations suggest an association between subtype F infection and heterosexual transmission (OR = 5.4; 95% CI 0.8756–34.6853) as reported previously [Rios et al., 2005]. The higher prevalence of women in the subtype F epidemic compared to subtype B epidemic coincides with the later introduction of the former in Brazil at the same time that the heterosexual transmission became more frequent. In contrast, the subtype B epidemic started in Brazil when the most common route of transmission of HIV was the homosexual contact followed by the intravenous drug users.

Five different patterns of subtypes B and F recombination between *gag* and *env* were identified, suggesting that HIV-1 inter-subtype recombination is occurring very frequently in this area. Among these recombinants, 18 samples shared the same breakpoint in *gag*. However, they presented different genotypes in *env*, meaning that (1) they can be the result of different recombination patterns or (2) they shared a common ancestry at some point. Other reports [Carr et al., 2001; Thomson et al., 2002; Quarleri et al., 2004; Sanabani et al., 2006] have described the co-circulation of BF recombinants with different genetic structures but presenting breakpoints at the same or at a very close position in the genome. This suggests the occurrence of hot spots for recombination in the HIV genome. In fact,

the initial region of *gag* seems to represent one of these hot spots since besides the recombinants identified in this cohort (position 1398 relative to HXB2), other unique BF recombinants [Carr et al., 2001; Thomson et al., 2002; De Sa Filho et al., 2005; Sanabani et al., 2006] as well as the CRF12 from Argentina (position 951) [Thomson et al., 2002], CRF28 and CRF29 (position 1322) [De Sa Filho et al., 2006] from Brazil presented crossover around these genome portions.

Two of the BF recombinants described previously (01BR226 and 02BR33) [Sanabani et al., 2006] grouped together with 11 of the new recombinants in both *gag* and *env* phylogenies (Fig. 3). This suggests that they shared common ancestries in both subtypes but as they present different breakpoints, they derived from distinct process of recombination. The same happened for two recombinants from Bahia (155 and 414), which fell in the same group of CRF28 and CRF 29 sequences. Eleven recombinants of this cohort sharing the same breakpoint in *gag* clustered in a monophyletic cluster apart from the CRF28/29 group pointing to different recombination origins. All recombinants of this group that have available *env* sequences, but one, also fell into a monophyletic group in both *env* tree and *gag*/*env* concatenated tree. These sequences were unrelated to CRF12. Taken together, these observations represent evidences that a new CRF is circulating in Bahia with a significant prevalence of 6.3%. Since the genetic distances of this group (6.6% and 15.3% in *gag* and *env*, respectively) are very similar to the intra-subtype distances, an early emergence and spread of these recombinants are also speculated. However, to confirm this hypothesis, studies involving their complete genome are needed. In addition, the *env* subtype B sequences of these recombinants clustered together with all subtype B sequences from Bahia described previously. This indicates that a main subtype B strain is present in the state and that it is unrelated to the subtype B strain that originated CRFs 28 and 29.

The presence of an F_{gag}/D_{env} recombinant HIV-1 in Brazil is reported in this study. Interestingly, the subtype D sequence (Fig. 3) did not cluster with the other subtype D sequences identified so far in the country [Guimarães et al., 2002; Couto-Fernandez et al., 2006] nor with the other subtype D sequences from Africa used as reference. However, its F part was related to other sequences from Brazil. Thus, this recombinant was likely generated locally around 10 years ago (diagnosis on 1998, Table I). These data also show that subtype D was introduced more than once in Brazil.

In conclusion, the HIV epidemic in Bahia is marked by a high genetic variability degree, with subtypes B, F, and D and their related recombinant forms being co-circulating. Subtype F was strongly related to heterosexually infected females. Similar to what has been observed for Brazilian southeastern region, variants with different recombination patterns between subtypes B and F, including a potentially new CRF, are present in the northeastern. An increased prevalence of these variants was detected in this geographic area. As the genetic and antigenic variability of HIV-1 represent a major challenge for the development of globally effective HIV vaccine, the understanding of the mechanisms driving these processes is crucial.

ACKNOWLEDGMENTS

We thank Mrs. Elisabeth Deliege Vasconcelos and Mr. Patrick Younan for editing and revising this manuscript, respectively.

REFERENCES

Barreto CC, Nishyia A, Araujo LV, Ferreira JE, Busch MP, Sabino EC. 2006. Trends in antiretroviral drug resistance and clade distributions among HIV-1-infected blood donors in Sao Paulo, Brazil. J Acquir Immune Defic Syndr 41:338–341.

Boletim Epidemiológico AIDS. Ano XVI n° 01. Available at http://www.aids.gov.br. Access on July 2007.

Buonaguro L, Tagliamonte M, Tornesello ML, Buonaguro FM. 2005. Evaluation of a modified version of heteroduplex mobility assay for rapid screening of HIV-1 isolates in epidemics characterized by mono/dual clade predominance. J Virol Methods 124:123–134.

Burke DS, McCutchan FE. 1996. Global distribution of human immunodeficiency virus type 1 clades. In: DeVita VT, Hellman L, Jr., Rosenberg SA, AIDS: Biology, diagnostics, treatment and prevention. 4th edition. New York: Lippencott-Raven Publishers, pp. 119–126.

Cabral VP, Cunha CB, Magalhaes EF, Pinto-Neto LF, Couto-Fernandez JC, Dietze R, Morgado MG, Ribeiro-Rodrigues R. 2006. Human immunodeficiency virus type-1 subtypes of infected patients in Espirito Santo, Brazil. Mem Inst Oswaldo Cruz 101: 881–885.

Carr JK, Avila M, Gomez Carrillo M, Salomon H, Hierholzer J, Watanaveeradej V, Pando MA, Negrete M, Russell KL, Sanchez J, Birx DL, Andrade R, Vinoles J, McCutchan FE. 2001. Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America. AIDS 15:F41–F47.

Casado G, Thomson MM, Sierra M, Najera R. 2005. Identification of a novel HIV-1 circulating ADG intersubtype recombinant form (CRF19_cpx) in Cuba. J Acquir Immune Defic Syndr 40:532–537.

Couto-Fernandez JC, Morgado MG, Bongertz V, Tanuri A, Andrade T, Brites C, Galvao-Castro B. 1999. HIV-1 subtyping in Salvador, Bahia, Brazil: A city with African sociodemographic characteristics. J Acquir Immune Defic Syndr 22:288–293.

Couto-Fernandez JC, Silva-de-Jesus C, Veloso VG, Rachid M, Gracie RS, Chequer-Fernandez SL, Oliveira SM, Arakaki-Sanchez D, Chequer PJ, Morgado MG. 2005. Human immunodeficiency virus type 1 (HIV-1) genotyping in Rio;1; de Janeiro, Brazil: Assessing subtype and drug-resistance associated mutations in HIV-1 infected individuals failing highly active antiretroviral therapy. Mem Inst Oswaldo Cruz 100:73–78.

Couto-Fernandez JC, Eyer-Silva WA, Guimarães ML, Chequer-Fernandez SL, Grinsztejn B, Delaporte E, Peeters M, Morgado MG. 2006. Phylogenetic analysis of Brazilian HIV type 1 subtype D strains: Tracing the origin of this subtype in Brazil. AIDS Res Hum Retroviruses 22:207–211.

De Sa Filho DJ, Sanabani S, Diaz RS, Munerato P, Brunstein A, Fusuma E, Sabino EC, Janini LM. 2005. Analysis of full-length human immunodeficiency virus type 1 genome reveals a variable spectrum of subtypes B and f recombinants in Sao Paulo, Brazil. AIDS Res Hum Retroviruses 21:145–151.

De Sa Filho DJ, Sucupira MC, Casiero MM, Sabino EC, Diaz RS, Janini LM. 2006. Identification of two HIV type 1 circulating recombinant forms in Brazil. AIDS Res Hum Retroviruses 22:1–13.

Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M, Rubsamen-Waigmann H, Mullins JI. 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: Analysis of HIV-1 *env* genes. Science 262:1257–1261.

Dourado I, Milroy CA, Mello MA, Ferraro GA, Castro-Lima Filho H, Guimarães ML, Morgado MG, Teixeira MG, Barreto ML, Galvão-Castro B. 2007. HIV-1 seroprevalence in the general population of Salvador, Bahia State, Northeast Brazil. Cad Saude Publica 23: 25–32.

Eyer-Silva WA, Couto-Fernandez JC, Morgado MG. 2007. Molecular epidemiology of HIV type 1 in inner Rio De Janeiro State, Brazil. AIDS Res Hum Retroviruses 23:303–308.

Guimarães ML, dos Santos Moreira A, Loureiro R, Galvão-Castro B, Morgado MG. Brazilian Network for HIV Isolation and Characterization. 2002. High frequency of recombinant genomes in HIV type 1 samples from Brazilian southeastern and southern regions. AIDS Res Hum Retroviruses 18:1261–1269.

Heyndrickx L, Janssens W, Zekeng L, Musonda R, Anagonou S, Van der Auwera G, Coppens S, Vereecken K, De Witte K, Van Rampelbergh R, Kahindo M, Morison L, McCutchan FE, Carr JK, Albert J, Essex M, Goudsmit J, Asjö B, Salminen M, Buvé A, van Der Groen G. 2000. Simplified strategy for detection of recombinant HIV-1 group M isolates by *gag*/*env* heteroduplex mobility assay. J Virol 74:363–370.

Kunanusont C, Foy HM, Kreiss JK, Rerks-Ngarm S, Phanuphak P, Raktham S, Pau CP, Young NL. 1995. HIV-1 subtypes and male-to-female transmission in Thailand. Lancet 345:1078–1083.

Mccutchan FE, Viputtigul K, De Souza MS, Carr JK, Markowitz LE, Buapunth P, McNeil JG, Robb ML, Nitayaphan S, Birx DL, Brown AE. 2000. Diversity of envelope glycoprotein from human immunodeficiency virus type 1 of recent seroconverters in Thailand. AIDS Res Hum Retroviruses 16:801–805.

Monteiro JP, Ferraro GA, Oliveira T, Goldani LZ, Kashima S, Alcantara LC, Morgado MG, Bou-Habib DC, Galvão-Castro B. 2007. Genetic and biologic characterization of HIV type 1 subtype C isolates from south Brazil. AIDS Res Hum Retroviruses 23:135–143.

Morgado MG, Guimaraes ML, Galvao-Castro B. 2002. HIV-1 polymorphism: A challenge for vaccine development—A review. Mem Inst Oswaldo Cruz 97:143–150.

Nicholas KB, Nicholas HB, Jr., Deerfield DW, II. 1997. GeneDoc: Analysis and visualization of genetic variation. EMBNET News 4:1–4.

Page RD. 1996. TreeView: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358.

Pedroso C, Queiroz AT, Alcântara LC, Drexler JF, Diaz RS, Weyll N, Brites C. 2007. High prevalence of primary antiretroviral resistance among HIV-1-infected adults and children in Bahia, a northeast state of Brazil. J Acquir Immune Defic Syndr 45:251–253.

Pereira GA, Stefani MM, Araujo Filho JA, Souza LC, Stefani GP, Martelli CM. 2004. Human immunodeficiency virus type 1 (HIV-1) and Mycobacterium leprae co-infection: HIV-1 subtypes and clinical, immunologic, and histopathologic profiles in a Brazilian cohort. Am J Trop Med Hyg 71:679–684.

Posada D, Krandall KA. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics 14:817–818.

Quarleri JF, Rubio A, Carobene M, Turk G, Vignoles M, Harrigan RP, Montaner JS, Salomón H, Gómez-Carrillo M. 2004. HIV type 1 BF recombinant strains exhibit different pol gene mosaic patterns: Descriptive analysis from 284 patients under treatment failure. AIDS Res Hum Retroviruses 20:1100–1107.

Rios M, Fernandez J, Jaramillo P, Paredes V, Sanchez JL, Laguna-Torres VA, Carr JK, Ramirez E. 2005. Molecular epidemiology of HIV type 1 in Chile: Differential geographic and transmission route distribution of B and F subtypes. AIDS Res Hum Retroviruses 21:835–840.

Salminen MO, Carr JK, Burke DS, McCutchan FE. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS Res Hum Retroviruses 11:1423–1425.

Sanabani S, Kleine Neto W, Kalmar EM, Diaz RS, Janini LM, Sabino EC. 2006. Analysis of the near full length genomes of HIV-1 subtypes B, F, and BF recombinant from a cohort of 14 patients in Sao Paulo, Brazil. Infect Genet Evol 6:368–377.

Swofford D. 1999. PAUP 4.0: 1999. Phylogenetic analysis using parsimony (and other methods), 4.0b2a. Sunderland, MA: Sinauer Associates, Inc.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEG A4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins D. 1997. The CLUSTAL_X Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882.

Thomson MM, Delgado E, Herrero I, Villahermosa ML, Vázquez-de Parga E, Cuevas MT, Carmona R, Medrano L, Pérez-Alvarez L, Cuevas L, Nájera R. 2002. Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Argentina revealed by analysis of near full-length genome sequences. J Gen Virol 83:107–119.

Thomson MM, Casado G, Posada D, Sierra M, Najera R. 2005. Identification of a novel HIV-1 complex circulating recombinant form (CRF18_cpx) of Central African origin in Cuba. AIDS 19:1155–1163.

Weniger BG, Takebe Y, Ou CY, Yamazaki S. 1994. The molecular epidemiology of HIV in Asia. [Retraction of Weniger BG, Takebe Y, Ou CY, Yamazaki S. In: AIDS 2001 Mar 9;15(4):545]. AIDS 8:S13–S28.