

Lower Prevalence of Human Immunodeficiency Virus Type 1 Brazilian Subtype B Found in Northeastern Brazil with Slower Progression to AIDS

Adriano Fernando Araujo,^{1,*} Carlos Brites,^{2,*} Joana Monteiro-Cunha,¹ Luciane Amorim Santos,¹ Bernardo Galvao-Castro,^{1,3} and Luiz Carlos Junior Alcantara^{1,3,4}

Abstract

Besides being extremely useful in measuring the level of HIV-1 diversity and prevalence in populations, the molecular analysis of genomic sequences provides crucial surveillance support and aids in the development of new therapies and effective vaccines. The present study focused on *gag* and *env* DNA and amino acid sequences that were generated from samples taken from 61 infected patients in the City of Salvador, Bahia, located in northeastern Brazil. In order to determine selective pressure and predict coreceptor usage, Bioinformatics tools were employed in phylogeny reconstruction. Fifty-six (91.8%) viruses were classified as belonging to subtype B, three (4.9%) from F1, and two (3.3%) from BF1 recombinants. Based on the characterization of the V3 region, the subtype B strains were represented by eight (18.2%) Brazilian variants (B'-GWGR), 20 (46.5%) European/EUA B variants (GPGR), and 15 (34.9%) GXGX variants. The mean time elapsed since diagnosis was 13 years among subtype B' and 9 years in subtype B. The mean dN/dS ratios from the GWGR, GPGR, and GXGX groups, when compared to an HXB2 reference, were 0.72, 0.77, and 0.67, respectively. Seventy-six percent of the viruses studied were predicted to use the CCR5 coreceptor for cell entry (R5 viruses), while 24% were predicted to use the CXCR4 or were classified as dual tropic viruses. The prevalence of subtypes B' and recombinant B/F1 was shown to be lower than findings from previous studies performed both in Brazil (B') and in Bahia (B/F1). The association between subtype B' and a lengthy period of time since diagnosis can be correlated with a slower disease progression in infected patients, when compared with those infected with subtype B.

Introduction

THE REMARKABLE GENOMIC VARIABILITY OF HIV is one of the main obstacles preventing immunologic infection control and impeding antiretroviral therapy efficacy, as well as vaccine development. This great variation has enabled virus classification and the study of viral genetic prevalence through phylogenetic analysis. HIV-1 group M variants, mainly responsible for the current HIV/AIDS pandemic, are divided into nine subtypes, circulating and unique recombinant forms (CRF and URF). These strains present distinctive geographic distribution worldwide, as well as in Brazil, with the highest number of HIV infections and greatest level of viral diversity found in Latin America.¹ The subtype B is the predominant genotype found in Brazil, followed by F1, C, and

recombinants B/F1.² Since 1993, a viral strain containing a GWGR signature motif at the tip of the V3 loop, named Brazilian subtype B (B'), has been detected in about half the samples analyzed^{3,4} and a slower AIDS progression has been associated with subtype B' when compared to subtype B.⁵ To gain entry into human cells, an HIV-1 particle interacts with a CD4 cellular receptor and a coreceptor, most commonly CCR5 or CXCR4. CCR5 usage (R5 virus) is associated with acute or asymptomatic patients, while CXCR4 (X4 virus) is associated with an elevated viral load, low CD4 T-cell counts, an increased disease progression rate, and a pronounced resistance to CCR5 inhibitors. HIV-1 tropism is frequently determined during clinical practice and antiretroviral therapy management. The aim of this study was to characterize the molecular diversity of HIV-1 in Salvador, the state capital of

¹Advanced Public Health Laboratory, Gonçalo Moniz Research Center, Oswaldo Cruz Foundation, Salvador, BA, Brazil.

²Federal University of Bahia, Salvador, BA, Brazil.

³Bahia School of Medicine and Public Health/Foundation for Development of Science, Salvador, BA, Brazil.

⁴Animal Models and Retroviral Vaccines Section, NCI/NIH, Bethesda, Maryland.

*These authors contributed equally to this work.

Bahia, which is located in the northeastern region of Brazil, and to compare genomic variability with the epidemiological and clinical characteristics of HIV-1-infected patients.

Materials and Methods

Sixty-one HIV-infected individuals, patients from the Professor Edgar Santos University Hospital (HUPES), located in the City of Salvador, Brazil, were recruited in 2006. All patients signed a letter of informed consent. Two to 5 ml of whole blood samples were collected and transported to the Advanced Laboratory of Public Health (LASP) at CPqGM/FIOCRUZ where they were stored at -20°C until use. Clinical data were also obtained from patient medical records. This study was approved by the CPqGM/FIOCRUZ Ethics Committee. DNA samples were extracted from 200 μl of blood using a QIAamp DNA kit (Qiagen Inc., Valencia, CA) in accordance with manufacturer's directions. Fragments of *gag* (positions 898 to 1968) and *env* (positions 6945 to 8183) genes, both relative to the HXB2 reference sequence, were PCR amplified using nested primers. The amplified DNA was purified using a purification kit (Qiagen Inc.) and sequenced using a BigDye Terminator v.3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Carlsbad, CA) and an automated ABI 3100 Genetic Analyzer (Applied Biosystems). Primers GAG2, G17, H1G777, and MZ14 were used for *gag* sequencing reactions, while primers ED31, MM4, ES7, and ED14 were used for the *env* region.

Electropherograms were obtained, analyzed, and exported in FASTA format using SeqScape v2.1.1 software (Applied Biosystems). Sequences were subjected to the BLAST search algorithm (<http://blast.ncbi.nlm>) to detect any evidence of contamination. Using a subtype reference set from the Los Alamos database (<http://hiv-web.lanl.gov>), an alignment was created using MUSCLE software⁶ and manually edited using GENEDOC software.⁷

Phylogenetic analyses were performed using PAUP* 4.0b10 software⁸ to generate neighbor-joining (NJ) and maximum likelihood (ML) trees using the GTR model of nucleotide substitution.⁹ Node reliability was assessed using bootstrap analysis (1000 replicates) and a likelihood-ratio test was used to calculate statistical support for tree branches. The trees were drawn using Figtree software (<http://tree.bio.ed.ac.uk>). Sequences that did not cluster within any pure subtype groups were submitted to recombination analysis using the Simplot software bootscanning tool¹⁰ and Genedoc software to determine the specific crossover point using visual inspection.

Coreceptor usage was predicted *in silico* based on V3-sequences and clinical data using the geno2pheno [coreceptor] online tool (<http://coreceptor.bioinf.mpiinf.mpg.de>) which is based on a statistical learning method that predicts CXCR4 usage by HIV-1.¹¹

Selective pressure was assessed in the 1155-bp envelope fragment using the SNAP¹² web-based tool (<http://www.hiv.lanl.gov>) which calculates nonsynonymous (dN) and synonymous (dS) substitution rates. Using these rates, the dN/dS ratio was calculated. Considering that a high dN/dS ratio suggests a trend towards positive selection,¹³ this ratio indicates the level of selective pressure through which the retrovirus isolates pass. Clinical, epidemiological, and molecular data were analyzed using PASW software, version 18.0 for Windows (SPSS Inc.). A statistically significant

difference was assumed when $p < 0.05$ and 95% CI did not cross zero.

Results

Out of 61 HIV-1 positive samples analyzed, 41 were obtained from males and 20 were from females (2:1 ratio). Among males, a mean age of 43.3 years was observed versus 39.1 years in females. Mean infection time was 9.94 years, mean viral load was 75,587.87 copies/ml and mean TCD4 cell counts were 421.61 cells/ml, while TCD8 cell counts were 1,262.24 cells/ml. Regarding ethnicity, 23% of the patients reported being of European or Latin descent, while 47.5% indicated Mixed-race and 29.5% African descent. No statistically significant differences regarding time of infection, viral load, or TCD4 and TCD8 cell counts were observed between ethnic groups. Out of 61 samples, 42 *gag* and 46 *env* gene sequences were amplified with 27 sequences amplified in both genes. Phylogenetic analysis of the *gag* gene showed 40 (95.2%) sequences clustered within the subtype B reference group, while two (4.8%) did not cluster within any pure subtypes and were further characterized as BF1 recombinants as revealed by recombination analysis. The BAS026 sequence showed two breakpoints (B/F1/B): one in positions 1357–1398 and another at positions 1482–1615 (relative to HXB2). The BAS096 sequence presents one breakpoint in positions 1187–1228 (relative to HXB2). In the *env* sequences, 43 (93.5%) were shown to cluster inside subtype B, and 3 (6.5%) inside F1 (Fig. 1). The two *gag* BF1 recombinants were subtyped as B in the *env* tree. With respect to the 27 samples that were characterized in both genes, 25 (92.6 %) belonged to subtype B, while only two (7.4 %) were found to be BF1/B. Considering all 61 samples, 56 (91.8%) were shown to be subtype B, three (4.9%) were F1, and two (3.3%) were BF1 recombinants. The 43 subtype B *env* sequences were translated into amino acid sequences. Based on V3 characterization, eight (18.2%) Brazilian (B'-GWGR), 20 (46.5%) European/EUA B (GPGR), and 15 (34.9%) GXGX variants were found. The mean pairwise distance was 0.11, 0.13, and 0.12 within the GWGR, GPGR, and GXGX groups, respectively. To test whether selective pressure was altered due to V3 loop substitution, the dN/dS ratio of each sequence was determined under comparison with an HXB2 reference sequence and the sequences were subsequently grouped in categories: GWGR, GPGR, and GXGX. The mean dN/dS ratio of all sequences was 0.725 when compared to HXB2. The mean GWGR, GPGR, and GXGX dN/dS group ratios were 0.72, 0.77, and 0.67, respectively, when compared to HXB2. Statistical testing revealed no statistically significant differences between the GWGR and GPGR groups ($p > 0.05$). A comparison between GPGR and GXGX group sequences showed a significant difference ($p = 0.018$), revealing a higher rate of positive selection in the GPGR group. We compared the clinical characteristics of individuals harboring B' (GWGR) and B (GPGR + GXGX) viruses. Table 1 shows that the mean time period since diagnosis was higher in the subtype B' group than in the subtype B group and that the mean age was also higher in the former group.

With respect to coreceptor usage, eleven (24%) out of the 46 V3 sequences were predicted to use the CXCR4 coreceptor (X4 virus). The mean age among subjects infected with X4 virus (32.4 years) was lower than the mean age among subjects infected with R5 virus (43.3 years) ($p < 0.05$).

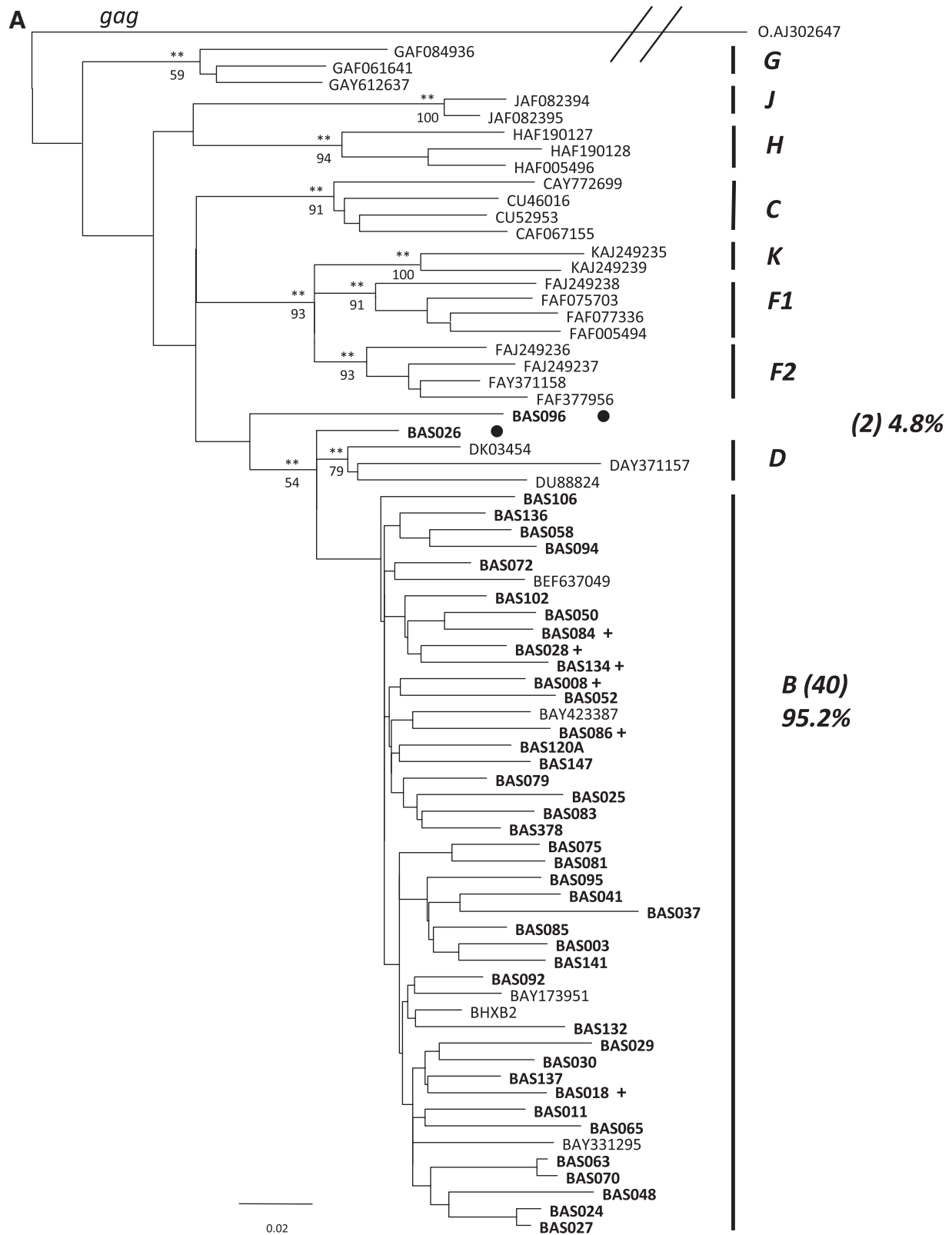


FIG. 1. NJ tree based on *gag* (A) and *env* (B) sequences showing phylogenetic relationships between HIV-1 samples and group M reference sequences. A sequence from group O was used as an outgroup sequence. The GTR+I+G nucleotide substitution model was used in both trees. Branches supported by the ML method are indicated by an asterisk (*) when significant ($p < 0.05$), and marked with a double asterisk (**) when highly significant ($p < 0.001$). Bootstrap values for 1000 replicates are shown in percentages. The HIV-1 sequences which were generated in the present study are indicated in **bold**. Subtype B' sequences are marked by a plus sign (+). Two sequences remained outside the subtype B and subtype D [indicated by a bullet point (●)].

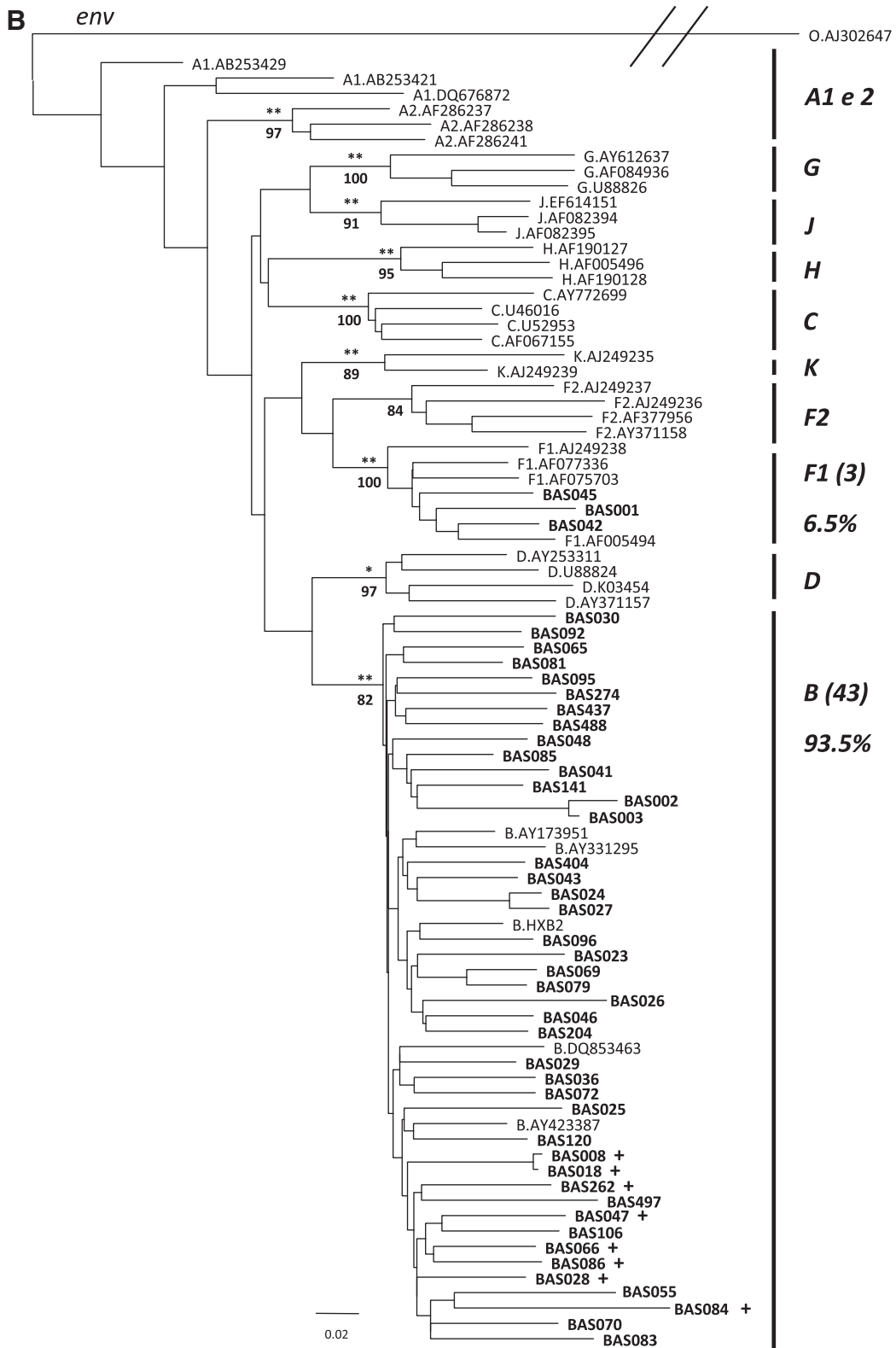


FIG. 1. (Continued).

Discussion

Molecular studies of HIV-1 are critical to furthering the understanding of mechanisms involved in AIDS pathogeny and to support the development of vaccines and efficacious

therapies. The number of HIV sequences from Northeastern Brazil remains scarce, as well as precise information associating HIV-1 diversity with clinical data. To this end, the authors studied samples and laboratory information from randomized patients in Salvador, the capital of the

TABLE 1. COMPARISON BETWEEN SUBTYPES B AND B' AND CLINICAL CHARACTERISTICS OF HIV-1-POSITIVE PATIENTS FROM SALVADOR, BRAZIL

	B×B'	N	Mean	Standard deviation	p*	95% (C.I.)	
						Lower	Upper
Age	B	34	39.09	12.983	0.057	-19.876	0.303
	B'	8	48.88	11.294			
Viral load	B	34	32202.35	81,611.235	0.857	-55584.779	66556.985
	B'	8	26716.25	48,895.262			
CD4	B	34	420.97	248.905	0.690	-232.868	155.559
	B'	8	459.63	222.836			
CD8	B	28	1439.32	1,391.531	0.538	-759.012	1428.798
	B'	7	1104.43	436.400			
Time since diagnosis	B	29	8.83	3.536	0.019	-7.087	-0.686
	B'	7	12.71	4.572			

*Test t for equality of means.

Northeastern Brazilian state of Bahia and the third most populous city in the country.

In the northeastern region of Brazil, as well as the country as a whole, HIV-1 subtype B remains predominant. However, in the past decade, several studies have reported on the increasing prevalence of other genotypes, notably BF recombinants.¹⁴ In the City of Salvador, a lower prevalence of B/F1 recombinant forms (3.3%) was observed in our samples when compared with recent studies showing a 10% and 13% prevalence of BF forms in the Brazilian Northeast.^{14,15} This may be explained by the fact that our sequencing involved two fragments within the *gag* and *env* genes, implying that other genomic regions where recombination could have occurred may have been overlooked.

Some of the sequences characterized in this study presented different clustering within the large subtype B group in *gag* and *env* trees. For instance, sequences BAS008 and BAS018 showed a close relationship in the *env* tree; however, their *gag* genes were not closely related in the phylogenetic analysis. These observations could be indicating that these samples derived from a common ancestral that went through different evolution process and subsequent intrasubtype recombination. Furthermore, these strains could be representing the occurrence of dual infection in one or both individuals. This study found the prevalence of subtype B' to be much lower than 50%, which was observed in the country's southeastern region.¹⁶ This leads us to suggest that Salvador has experienced different introduction(s) and founder effect(s) from the Brazilian Southeast. In addition, no phylogenetic distinctions were observed between subtypes B' and B, findings consistent with previous reports.¹⁶ However, this study demonstrated that mean age (49 years) and time period elapsed since diagnosis (13 years) in subtype B'-infected patients was higher than mean age (39 years) and time period elapsed since diagnosis (9 years) for subtype B-infected individuals. This could be related to an increased replication rate and accelerated disease progression characteristic of subtype B sequences.⁵ In fact, the GPGR sequences exhibited a slightly higher mean pairwise distance which may explain the clinical difference between subtype B' and subtype B.

The results presented in this study demonstrate a lower level of genetic diversity and, specifically, a lower prevalence of BF recombinant forms than those previously found

in Bahia and in the Brazilian Northeast region.^{14,15} This may be related to the presence of recombination points in genomic regions other than those analyzed by this study, and, therefore, further studies involving the full genome sequencing of HIV isolates from this geographic region could contribute to a better understanding of this region's HIV epidemiology.

Acknowledgments

The authors are grateful to the individuals who donated blood for the purposes of this study, to the FIOCRUZ-PDTIS sequence platforms, Mr. Augusto Santana and Mrs. Maurina Alcantara from HUPES for providing access to patient biomedical records, and to Mrs. Elisabeth Deliege Vasconcelos for editing and revising this manuscript.

Sequence Data

The new sequences in this study were reported to GenBank under the accession numbers GU595197-281 and GU722093-95.

Author Disclosure Statement

No competing financial interests exist.

References

- Bastos FI, Cáceres C, Galvão J, Veras MA, and Castilho EA: AIDS in Latin America: Assessing the current status of the epidemic and the ongoing response. *Int J Epidemiol* 2008;37:729-737.
- Morgado MG, Guimarães ML, and Galvão-Castro B: HIV-1 polymorphism: A challenge for vaccine development. A review. *Mem Inst Oswaldo Cruz* 2002;97:143-150.
- Potts KE, Kalish ML, Lott T, *et al.*: Genetic heterogeneity of the V3 region of the HIV-1 envelope glycoprotein in Brazil. Brazilian Collaborative AIDS Research Group. *AIDS* 1993;7:1191-1197.
- Morgado MG, Sabino EC, Sphaer EG, *et al.*: V3 region polymorphisms in HIV-1 from Brazil: Prevalence of subtype B strains divergent from North American/European prototypes and detection of subtype F. *AIDS Res Hum Retroviruses* 1994;10:569-576.

5. De Brito A, Komninakis SC, Novoa P, *et al.*: Women infected with HIV type 1 Brazilian variant, subtype B (B'-GWGR motif) have slower progression to AIDS, compared with patients infected with subtype B (B-GPGR motif). *Clin Infect Dis* 2006;43:1476-1481.
6. Edgar RC. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 2004;5:113-132.
7. Nicholas KB, Nicholas Jr HB, and Deerfield II DW: GeneDoc: Analysis and visualization of genetic variation. *Embnet News* 1997;4:1-4.
8. Swofford D. PAUP 4.0: Phylogenetic analysis using parsimony (and other methods), 8 4.0b2a. Sunderland, MA, USA: Sinauer Associates, Inc. 1999.
9. Posada D and Krandall KA: MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 1998;14:817-818.
10. Salminen MO, Carr JK, Burke DS, and McCutchan FE: Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retrovir* 1995;11:1423-1425.
11. Sing T, Low AJ, Beerenwinkel N, *et al.*: Predicting HIV coreceptor usage based on genetic and clinical covariates. *Antiviral Therapy* 2007;12:1097-1106.
12. Korber B: HIV Signature and Sequence Variation Analysis. In: *Computational Analysis of HIV Molecular Sequences*. (Allen G. Rodrigo and Gerald H. Learn, eds). Dordrecht, Netherlands: Kluwer Academic Publishers. 2000, pp. 55-72.
13. Mindell DP: Positive selection and rates of evolution in immunodeficiency viruses from humans and chimpanzees. *Proc Natl Acad Sci USA* 1996;93:3284-3288.
14. Monteiro JP, Alcantara LC, de Oliveira T, *et al.*: Genetic variability of human immunodeficiency virus-1 in Bahia state, Northeast, Brazil: High diversity of HIV genotypes. *J Med Virol* 2009;81:391-399.
15. Brennan CA, Brites C, Bodelle P, *et al.*: HIV-1 strains identified in Brazilian blood donors: Significant prevalence of B/F1 recombinants. *AIDS Res Hum Retroviruses* 2007;11:1434-1441.
16. Diaz RS, Leal E, Sanabani S, *et al.*: Selective regimes and evolutionary rates of HIV-1 subtype B V3 variants in the Brazilian epidemic. *Virology* 2008;38:184-193.

Address correspondence to:

Luiz CJ Alcantara
Building 41, Room C-303
National Cancer Institute/NIH
Bethesda, MD 20892-5065

E-mail: alcantaralc@mail.nih.gov