Distribution of human immunodeficiency virus type 1 subtypes in the State of Amazonas, Brazil, and subtype C identification

L.K.H. Cunha, S. Kashima, M.F.C. Amarante, R. Haddad, E.S. Rodrigues, K.L.T. Silva, T.A. Lima, D.B. Castro, F.C. Brito, E.G. Almeida, D.T. Covas and A. Malheiro
Distribution of human immunodeficiency virus type 1 subtypes in the State of Amazonas, Brazil, and subtype C identification

L.K.H. Cunha1, S. Kashima3, M.F.C. Amarante3, R. Haddad3, E.S. Rodrigues3, K.L.T. Silva2, T.A. Lima2, D.B. Castro2, F.C. Brito2, E.G. Almeida2, D.T. Covas3 and A. Malheiro1,2

1Departamento de Parasitologia, Universidade Federal do Amazonas, Manaus, AM, Brasil
2Diretoria de Ensino e Pesquisa, Fundação de Hematologia e Hemoterapia do Amazonas, Manaus, AM, Brasil
3Laboratório de Biologia Molecular, Hemocentro de Ribeirão Preto, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil

Abstract

Few studies have reported the molecular epidemiological characterization of HIV-1 in the Northern region of Brazil. The present study reports the molecular and epidemiological characterization of 31 HIV-1 isolates from blood donors from the State of Amazonas who donated blood between April 2006 and March 2007. Serum/plasma samples from all donors were screened for HIV antibodies by ELISA and the results confirmed by Western blot analysis. Genomic DNA was extracted from the buffy coat using the Super Quik-Gene-DNA Isolation kit. Nested PCR was performed on the env, gag, and pol regions of HIV-1 using the Gene Amp PCR System 9700. Sequencing reactions were performed using the inner PCR primers and the DYEnamic™ ET Dye Terminator Kit, and phylogenetic analysis was performed using the gag, pol, and env gene sequences. We collected samples from 31 blood donors who tested positive for HIV-1 in confirmatory experiments. The male:female ratio of blood donors was 3.4:1, and the mean age was 32.4 years (range: 19 to 61 years). Phylogenetic analysis showed that subtype B is the most prevalent among Northern Brazilian HIV-1-seropositive blood donors. One HIV-1 subtype C and one circulating recombinant form (CRF_BF) of HIV-1 were identified in the State of Amazonas. This is the first study showing the occurrence of a possible "homogenous" subtype C in this region of Brazil. This finding could contribute to a better characterization of the HIV-1 strains that circulate in the country.

Key words: HIV-1; Subtypes; Phylogenetic analysis; Blood donors; Molecular and epidemiological characterization

Introduction

Two major reasons for the extensive genetic diversity of HIV-1 are high levels of viral replication and error-prone reverse transcription, which incorporates mutations into the viral genome. Investigating the genetic diversity of HIV-1 can lead to further understanding of HIV-1 evolution and dissemination and vaccine development.

The HIV-1 genome is composed of approximately 9.0 kb distributed among structural (gag, env, and pol), trans-activation (tat and rev) and accessory (nef, vpu, vif, and vpr) genes. Based on phylogenetic analyses of env and gag gene sequences, HIV-1 is classified into three main groups: the major (M), outlier (O), and new (N) groups (1-4). Nine genetic subtypes have been identified in the M group: A-D, F-H, J, and K; at least 34 circulating recombinant forms (CRFs) have been recognized (5-7). These genetic subtypes may have an impact on drug susceptibility, the emergence of new drug resistance mutations, and the performance of laboratory tests for diagnostics and measurements of viral loads (8).

About one third of all people with HIV-1 in Latin America reside in Brazilian territory (9), where HIV-1 subtype B was originally predominant. This subtype was introduced in Brazil by North American individuals and established itself in the risk group of homosexual men (10). Previous studies have reported a variable prevalence of HIV-1 subtypes in Brazil, including subtypes F1, C, and D (11-13) and the CRFs...
(CRF_BC, BF) (14,15). Subtype F1 was identified primarily in female prostitutes and drug users, and was introduced in Brazil in the early 1980s (10). Although most HIV-1 infections in Brazil are subtype B, the southern region of the country represents the main subtype C focus (16). Isolated cases of this subtype have also been reported in the States of Espírito Santo, Goiás, and Rio de Janeiro (17-20).

The Northern region of Brazil contains 7 States (Rondônia, Acre, Amazonas, Roraima, Pará, Amapá, and Tocantins). Amazonas is located in the central region, comprises the largest territorial area of Brazil and shares borders with Venezuela, Colombia, and Peru. According to Boletim Epidemiológico DST/AIDS (2006), the Northern region reports approximately 13,000 AIDS cases, which corresponds to 3% of all cases in Brazil. The State of Amazonas has around 5000 notified AIDS cases, and Manaus (the capital of Amazonas) is responsible for 90% of these cases. However, few genotyping studies of HIV-1-infected people in the Northern region of Brazil have been reported. Recently, a study describing the presence of subtypes D and C in the cities of Belém (Pará) and Macapá (Amapá) has been reported (21), but the epidemiology of HIV in the State of Amazonas is unknown. The only study that describes HIV-1 subtypes in Manaus, the largest community within the central rain forest of Brazil’s Amazon Basin, showed that subtypes B and F are represented equally (22).

In the present study, we performed the molecular and epidemiological characterization of 31 HIV-1 isolates from the Northern region of Brazil. We also reported, for the first time, the presence of subtype C in the State of Amazonas.

Material and Methods

Patient samples

The samples were obtained from HIV-1-seropositive blood donors at Fundação Hemo- centro de Manaus (HEMOAM), AM, Brazil, and presented positive serology to HIV-1. The study was approved by the Institutional Ethics Committee and all blood donors signed an informed consent form and were informed about the procedures of sample collection and analysis. Between April 2006 and March 2007, 10-mL intravenous blood samples were collected from each individual into tubes containing the anticoagulant EDTA. Serum/plasma samples from all donors were screened for HIV antibodies by ELISA AxSYM HIV I/II gO (Abbott, Germany) and combined HIV Ag/Ab (Murex Biotech, UK). Confirmatory screening tests were performed by Western blot analysis using HIV Blot 2.2 (Genelabs Diagnostics, Singapore).

**Proviral DNA extraction**

DNA was extracted from the buffy coat (ELA-positive samples) using the Super Quik-Gene-DNA Isolation kit (Promega, USA) according to manufacturer instructions.

**Polymerase chain reaction (PCR)**

Nested PCR of HIV-1 was performed on the env, gag, and pol regions using a Gene Amp PCR System 9700 (Applied Biosystems, USA). To amplify a fragment of the genes, the first round of PCR was performed on samples using the primers ED5 and ED12 for the env gene, H1G777 and H1P202 or C1 and C2 for the gag gene, and K1 and K2 for the pol gene (18). A second round of PCR was carried out using the nested primers ED31 and ED33 for env, H11584 and G17 or SK and C2 for gag, and F1 and F2 for pol amplification. The primer sequences of HIV-1 genomic regions are presented in Table 1. Reactions were performed in 50-µL mixtures containing 500 ng DNA, 1.0 U Taq DNA polymerase, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 0.8 mM each deoxynucleotide triphosphate (dNTPs), and 10 pmol of each primer. PCR cycling conditions for the first and second rounds consisted of one cycle at 94°C for 5 min, 35 cycles at 94°C for 90 s, 55°C for 90 s, and 72°C for 30 s. Five microliters of the initially amplified product

| Region | Description | 5’-3’ |
|---------|-------------|-------|
| **Gag** | 1st round | TCA CCT AGA ACTTTG AATTGCA TGG G |
| | H1G777 | TCA CCT AGA ACTTTG AATTGCA TGG G |
| | H1P202 | CTA ATAC TAG ATG CAT CTG CTG CTG T |
| | C1 | GGT CAG CCAAATA TTT CCCTAT A |
| | C2 | GGA AAT GTG GAA AGG AAG GAC A |
| | 2nd round | AAA GAT GGA TAAA TCC TGG G |
| | G17 | TCC ACA TTT CCAAACA GCC CTT TTT |
| | SK | ATA ATC CAC CTA TCC CAG TAG GAG A |
| | C2 | GGA AAT GTG GAA AGG AAG GAC A |
| **Pol** | 1st round | CAG AGC CAA CAG CCC CAC CA |
| | K1 | TTT CCC CAC TAA CTT CTG TAT GTC ATT GAC A |
| | K2 | GGT GAC TCA GAT TGG TTG CAC |
| | F1 | GTA TGCT AAT CAG CAG TCC AGC |
| | F2 | GTA TGT CAT TGA CAG TCC AGC |
| **Env** | 1st round | ATGGGA TCAAAG CCA AAACA GCCTATG |
| | ED5 | ATGGGA TCAAAG CCA AAACA GCCTATG |
| | ED12 | AGGTGGCT TCC TGC TGC TGC TCC CAA GAA CCC AAG |
| | ED31 | CTA CAG CCA TTA CAC AGG CCT GTC CAA AG |
| | ED33 | TTA CAG TAG AAA AAT TCC CCT C |
(first round) was used in the second round. The amplified products were analyzed by 2% agarose gel electrophoresis followed by ethidium bromide staining.

**Nucleotide sequencing and phylogenetic analysis**

Amplified products (gag, pol, and env) were purified using the Wizard PCR Prep DNA Purification System Kit (Promega). Sequencing reactions were performed using the inner PCR primers and the DYEnamic™ ET Dye Terminator Kit according to manufacturer instructions, and products were sequenced using the MegaBace 1000 DNA Sequencing System. Electropherograms were analyzed with ChromasPro Version 1.41 (Technelysium Pty Ltda., Australia).

The HIV-1 sequences of Brazilian isolates (gag, pol, and env) were aligned and edited using the BioEdit 7.0.5.3 program. The sequences were compared to those in the Los Alamos database, which includes 40 prototypes for the genotypes A-D, F-H, J, and K, 14 CRFs and the out group sequence (O group). Neighbor joining (NJ) trees were constructed using the Phylip package (version 3.6) and PAUP* 4.0b10 program (Sinauer Associates, USA). The NJ trees were constructed using the (GTR+I+G) and (TMV+I+G) nucleotide substitution models, as selected by the Modeltest 3.06 program. The reliability of the NJ trees was assessed by analyzing 1000 bootstrap replicates. The likelihood ratio test was used to calculate statistical support (reported as P values) for the branches. Trees were drawn with the TreeView X 0.5.0 program. Sequences obtained in this study are available in GenBank under accession numbers FJ011469-FJ011532.

Recombination analysis was performed using SimPlot 3.5.1, which evaluates the differences between phylogenetic trees at each fragment (window) along with the alignment. Bootscanning was performed with alignment (available at http://lasp.cpqgm.fiocruz.br/retrovirusHIV.html). The gag, pol, and env regions were concatenated to get the query sequence, gaps were deleted to obtain a continuous sequence, and the sequences were then analyzed by multiple alignment.

**Results**

**Clinical features of HIV-1-infected patients**

We collected samples from a total of 107 HIV-1-seropositive blood donors who had one positive serology test. Of these 107 donors, only 31 were confirmed as HIV-1 positive in diagnostic tests. The majority of the HIV-1 individuals were males, the male:female ratio was 3.4:1, and the mean age was 32.4 years (range: 19 to 61 years; Table 2). Ten repeat blood donors were identified (32.3%), and 6 of them seroconverted for up to 20 months. The seroconversion time comprises the period between the last liberated donation and the donation that presented HIV-1 reactivity. The blood donors were asymptomatic at the time of donation and reported not to be under antiretroviral treatment (data not shown).

**Phylogenetic analysis**

The DNAs of 31 HIV-1-positive individuals were evaluated. We analyzed 462, 706, and 498 bp of the gag, pol, and env regions, respectively (Table 1). The NJ trees constructed from the three regions show that subtype B is the most prevalent among Brazilian patients (Figures 1, 2, and 3).

The phylogenetic trees revealed two subtype C samples. Only the env and pol regions were characterized in one sample (AM090) because the gag region did not amplify in PCR (Table 3). In the other sample (AM107), all regions were characterized. These samples were submitted to bootscan analysis, and the subtype C profiles were confirmed (data not shown).

Subtype C has not been previously described in the Amazonas State, but has been isolated in other regions of Brazil. Analyses of similarity were performed to compare the env sequences found in the Southern region of Brazil (DQ358770, DQ358766, DQ358763, DQ358761, DQ358757, DQ358756, and U52953) and env sequences from different geographic regions, including Venezuela (AY456916 and AY456917), South Africa (AY772699), India (AF087155), and Ethiopia (U46016) (Table 4). This genomic region was selected for similarity analysis because it is the most variable region of HIV-1. These analyses showed that Amazonian C subtype isolates show high similarity to sequences from the Southern region, suggesting that HIV-1 subtype C may have disseminated to the North following a South to North gradient pattern.

The partial sequences of gag, pol, and env genes were obtained for 18 samples, while for 10 samples, only two regions (gag and env, gag and pol, or pol and env) were obtained. For three samples, only one region (gag, pol, or env) was obtained (Table 3). In agreement with the topology of the phylogenetic trees built for each gene, 17 samples were classified as subtype B and one as subtype C for the three genes. However, one sample was classified as subtype B for the gag and pol genes and subtype F for the env gene, suggesting the presence of one circulating recombinant form.
The 828-bp concatenated sequence of probable CRF (sample AM047) was submitted to bootscanning using SimPlot with the following parameters: window size (300 bp), step (20 bp), algorithm (NJ), distance model (Kimura 2-parameter), and replicate number (100). The reference group/subtypes used for bootscanning were the M (A-D, F-H, J, K) and O groups. The map generated by the SimPlot software showed that the AM047 sample presented a breakpoint (subtypes B and F) at position 525. At 230, the sequence is characterized as subtype B (95% reliability), while at position 720 it is characterized as subtype F (99% reliability; data not shown).

**Discussion**

Several HIV-1 subtypes, sub-subtypes, circulating re-
combinant forms, and unique recombinant forms (URFs) have been described extensively worldwide (23). The HIV-1-positive prevalence among Brazilians is approximately 0.5%, ranging from 0.3 to 1.6% since 2000 (24). Most of the molecular and epidemiological characterization of HIV-1 in Brazil has concentrated in the Southeast, where HIV-1 subtype B is prevalent. Other Southern regions in the country have reported a predominance of the C subtype; however, the non-B and non-F subtypes may be neglected in other regions of the country (25-27). A survey of blood donors between 1996 and 2004 (over 400,000 donations) showed that the HIV-1/2 seroprevalence is about 0.2% in the Northern region of Brazil (HEMOAM Foundation, Blood Bank of Amazon State, unpublished data). In agreement with other regions, the present study showed a higher prevalence of subtypes B and F among recent blood donors and the presence of one recombinant B/F form (AM047: B<sub>pol</sub>, B<sub>gag</sub>, F<sub>env</sub>). Moreover, the present study also identified

Figure 2. Rooted NJ tree of 23 HIV-1 strains based upon a 706-bp fragment of the pol region. The bootstrap values (above 50% and using 1000 bootstrap samples) on the branches represent the percentage of trees for which the sequences at one end of the branch form groups. The OAY169812 strain is used as an outgroup. The Brazilian isolates are in bold. The tree shows that 91.3% of samples are subtype B, and 8.7% are subtype C.
two C subtype isolates (AM090: C\text{pol}, C\text{env}; AM107: C\text{pol}, C\text{gag}, C\text{env}). Subtype C is the most prevalent subtype of HIV-1 worldwide and accounts for more than half of infections. Recently, one study performed in the States of Pará and Amapá in Northern Brazil showed CRFs containing subtype C (C\text{env}, B\text{pro}) (21). Our data indicate, for the first time, the possible presence of a “homogenous” subtype C virus in the Amazon Basin. Complete genome sequencing of AM090 and AM107 isolates is required to confirm this hypothesis.

Figure 3. Rooted NJ tree of 28 HIV-1 strains based upon a 498-bp fragment of the env region. The bootstrap values (above 50% and using 1000 bootstrap samples) on the branches represent the percentage of trees for which the sequences at one end of the branch form groups. The OAY169812 strain is used as an outgroup. The Brazilian isolates are in bold. The tree shows that 85.7% of samples are subtype B, 7.1% are subtype C, and 7.1% are subtype F.
Two hypotheses may explain the route of HIV-1 subtype C transmission. First, although subtype B is the most prevalent subtype in Brazil, subtype C is present mainly in the Southern region (28). This subtype was previously reported to be found in Argentina, Uruguay, Paraguay (29), and other countries in South America. Recently, Jones et al. (30) reported that subtype C was introduced into South America from Argentina, which acquired African strains. However, other studies support the view that Brazil is the dispersion center of subtype C in South America. The exact origin of subtype C cannot be determined due to the lack of epidemiological and historical data, sample problems and doubts regarding phylogeny (30). Subtype C is concentrated in the Southern region of this continent and other states of Brazil, including Espírito Santo (Southeast region) and Goiás (Center-West region) (17-19). Our analysis shows that the Northern subtype C isolates are more similar to the Southern subtype C samples than in other regions, supporting the idea that subtype C has been disseminated in Brazil following a South to North gradient pattern.

In 2005, subtypes C and B/C were confirmed in Venezuelan patients who took trips to Africa and Europe. One patient (subtype C) was infected in South Africa, and another patient (CRF B/C) had high sexual risk behavior and a history of trips to Portugal and South Africa (31). These subtypes are also described in other countries near the Amazon Basin, including Peru and Ecuador (32). The second hypothesis is that subtype C was recently introduced into the Northern region of Brazil through countries that border this region. In the present study, it was not possible to identify the transmission pattern of HIV-1 in the State of Amazonas.

In Brazil, the subtypes B, F1 and C coexist in high risk groups and the prevalence differs only by geographical location (10,33). The coexistence of several subtypes in a region leads to the emergence of URFs and CRFs. While CRFs B/F1 are common in areas where subtypes B and F1 coexist, in the South of the country CRFs B/C are emerging (19) due to the presence of these two subtypes. The CRF B/F identified in the present study in Manaus corroborates the theory that these forms are recombining in different areas, increasing the genetic variability of HIV-1 in Brazil.

The present study reports one of the first global views of the molecular epidemiology of HIV-1 among Northern Brazilian isolates. We demonstrated by sequence analysis that HIV-1-infected individuals harbor the subtypes B, F and C, and different and/or complete genomic regions of HIV-1 should be analyzed in order to genetically characterize the Brazilian strains in this region. These studies will contribute to an understanding of the divergence of Brazilian strains of HIV-1 and the transmission patterns of this infection in our country. HIV-1 molecular epidemiology studies should further our understanding of the biological properties of HIV-1 and the mutations related to antiretroviral resistance and vaccine design.

**Acknowledgments**

We thank Rochele Azevedo, Andréia Monteiro Tarragô and Adriana Aparecida Marques for technical assistance. Research supported by CNPq, Fundação de Amparo à Pesquisa do Estado do Amazonas, Center for Cell-Therapy (CTC)/Fundação Hemocentro de Ribeirão Preto, and FAPESP.
References

1. Louwagie J, McCutchan F, Mascola J, Eddy G, Fransen K, Peeters M, et al. Genetic subtypes of HIV-1. *AIDS Res Hum Retroviruses* 1993; 9: S147-S150.

2. Louwagie J, McCutchan FE, Peeters M, Brennan TP, Sanders-Buell E, Eddy GA, et al. Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. *AIDS* 1993; 7: 769-780.

3. McCutchan F, Louwagie J, Eddy G, Van Der Groen G, Piot P, Myers G, et al. Identification of multiple genetic subtypes of HIV-1 and evidence for geographic dispersal. *J Acquir Immune Defic Syndr* 1993; 6: 685.

4. Simon F, Mauchere P, Roques P, Loussert-Ajaka I, Muller-Truttin MC, Saragosti S, et al. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med* 1998; 4: 1032-1037.

5. Takebe Y, Kusagawa S, Motomura K. Molecular epidemiology of HIV: tracking AIDS pandemic. *Pediatr Int* 2004; 46: 236-244.

6. Thomson MM, Najera R. Molecular epidemiology of HIV-1 variants in the global AIDS pandemic: an update. *AIDS Rev* 2005; 7: 210-224.

7. Los Alamos Laboratory. The circulating recombinant forms (CRFs). http://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html.

8. Koch WH, Sullivan PS, Roberts C, Francis K, Downing R, Mastro TD, et al. Evaluation of United States-licensed human immunodeficiency virus immunoassays for detection of group M viral variants. *J Clin Microbiol* 2001; 39: 1017-1020.

9. Joint United Nations Programme On HIV/AIDS (UNAIDS) and World Health Organization (WHO). Latin America: AIDS epidemic update: Regional summary. http://data.unaids.org/pub/Report/2008/jc1530_epibriefs_latinamerica_en.pdf. Accessed July 25, 2008.

10. Bello G, Guimaraes ML, Morgado MG. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS* 2006; 20: 763-768.

11. Sabino EC, Shpaer EG, Morgado MG, Corber BT, Diaz RS, Bongertz V, et al. Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals from Brazil. *J Virol* 1994; 68: 6340-6346.

12. Guimaraes ML, dos Santos MA, Loureiro R, Galvao-Castro B, Morgado MG. High frequency of recombinant genomes in HIV-1 type 1 samples from Brazilian southeastern and southern regions. *AIDS Res Hum Retroviruses* 2002; 18: 1261-1269.

13. Couto-Fernandez JC, Eyer-Silva WA, Guimaraes ML, Chequer-Fernandez SL, Grinszttein B, Delaporte E, et al. Phylogenetic analysis of Brazilian HIV type 1 subtype D strains: tracing the origin of this subtype in Brazil. *AIDS Res Hum Retroviruses* 2006; 22: 207-211.

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

15. Sa Filho DJ, Sanabani S, Diaz RS, Munerato P, Brunstein A, Fusuma E, et al. Analysis of full-length human immunodefi-
29. Carrion G, Eyzaguirre L, Montano SM, Laguna-Torres V, Serra M, Aguayo N, et al. Documentation of subtype C HIV type 1 strains in Argentina, Paraguay, and Uruguay. *AIDS Res Hum Retroviruses* 2004; 20: 1022-1025.

30. Jones LR, Dilernia DA, Manrique JM, Moretti F, Salomon H, Gomez-Carrillo M. In-depth analysis of the origins of HIV type 1 subtype C in South America. *AIDS Res Hum Retroviruses* 2009; 25: 951-959.

31. Castro E, Moreno M, Deibis L, de Perez G, Salmen S, Berrueta L. Trends of HIV-1 molecular epidemiology in Venezuela: introduction of subtype C and identification of a novel B/C mosaic genome. *J Clin Virol* 2005; 32: 257-258.

32. Montano SM, Sanchez JL, Laguna-Torres A, Cuchi P, Avila MM, Weissenbacher M, et al. Prevalences, genotypes, and risk factors for HIV transmission in South America. *J Acquir Immune Defic Syndr* 2005; 40: 57-64.

33. Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brigido L, et al. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. *AIDS* 2003; 17: 1063-1069.