JAIDS Journal of Acquired Immune Deficiency Syndromes
23:327–331 © 2000 Lippincott Williams & Wilkins, Inc., Philadelphia

The HIV Epidemic in the Amazon Basin Is Driven by Prototypic and Recombinant HIV-1 Subtypes B and F

*Ana C. P. Vicente, *Koko Otsuki, †Nelson B. Silva, †Marcia C. Castilho, †Flavio S. Barros, ‡Danuta Pieniazek, ‡Dale Hu, ‡Mark A. Rayfield, §Gustavo Bretas, and ‖Amilcar Tanuri

**Genetics Department, Instituto Oswaldo Cruz, Rio de Janeiro, Brazil; †Instituto de Medicina Tropical, Manaus, Brazil; ‡Division of AIDS, STD, and TB, Laboratory Research, National Center for Infectious Diseases, and Division of HIV/AIDS Prevention Surveillance and Epidemiology, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia, U.S.A.; §Universidade Estadual do Rio de Janeiro, Rio de Janeiro, Brazil; ‖Department of Genetic, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*

Summary: This paper describes genetic subtypes of HIV-1 found in blood samples from 31 HIV-1–infected people who visited the Counseling and Testing AIDS Center of Instituto de Medicina Tropical in Manaus, Brazil. Manaus, the main city in Brazil's Amazon Basin, is also the closest urban connection for more than 100,000 Indians living in the rain forests of this region. Although to date there is no evidence of increased incidence of HIV-1 infection among the indigenous population, our understanding of both the prevalence and nature of the epidemic in the region as a whole is limited. From the 31 samples analyzed by C2V3 sequencing, we found almost equal proportions of HIV-1 strains belonging to subtype B ($n = 16$; 51.6%) and subtype F ($n = 15$; 48.4%), a finding that differs from results from previous studies conducted in urban areas of southeastern Brazil. We also observed the presence of the GWGR amino-acid sequence in the critical tetra-peptide crown of the env V3 loop in the HIV-1 subtype B samples analyzed. Among these samples, we also found 14 mosaic genomes (45.16%) in which different combinations of subtypes B, C, and F were identified between the p24 *gag*, *pro*, and *env* regions. Our data support the hypothesis that the Amazonian HIV-1 infections linked to the urban epidemic in southeastern Brazil. The genetic diversity and the prevalence of mosaic genomes among the isolates in our study confirm an integral role of recombination in the complex Brazilian epidemic.
Key Words: HIV-1—Recombinant—Amazon basin.

The HIV/AIDS pandemic is characterized by high genetic diversity of its causative agent, HIV-1, and a complex pattern of microepidemics occurring in different regions of the world. On the basis of phylogenetic analysis, HIV-1 can be classified into three distinct groups: M (for "major"), O (for "outlier"), and N (for "new"). The group M variants have been further subdivided into 10 different subtypes (A–J) based on *gag* and *env* sequence analysis (1,2). Although HIV-1 subtypes A, B, C, D, and E predominate globally, subtype B viruses are disproportionately concentrated in Europe, Asia, and the Americas. The less prevalent group O variants represent highly divergent, genetically distinct HIV-1 strains. These variants have been isolated from persons with links to West Central Africa and are limited in their geographic distribution. Recently, HIV-1 group N has been proposed to describe a small cluster of African isolates that are genetically related to the chimpanzee simian immunodeficiency virus (SIV_{cpz}) (3).

As the largest country in South America, Brazil also has been most heavily affected by the HIV/AIDS epi-

Address correspondence and reprint requests to Amilcar Tanuri, Laboratorio de Virologia Molecular, Dept. of Genetic-IB, CCS Bloco A, Universidade Federal do Rio de Janeiro, Ilha do Fundao, Rio de Janeiro, RJ 21943-990, Brazil: email: lavimoan@hotmail.com.

Use of trade names is for identification only and does not constitute endorsement by the U.S. Department of Health and Human Services, the Public Health Service, or the Centers for Disease Control and Prevention.

Manuscript received June 24, 1999; accepted November 24, 1999.

demic. The Brazilian epidemic is characterized by the presence of multiple HIV-1 group M subtypes, primarily subtype B and the less prevalent subtype F, but also subtypes C and D (4–6). Subtype B viruses are represented by two variants distinguished by their V3 loop amino-acid sequence. The tetrameric sequence at the tip of the loop in almost half of Brazilian subtype B viruses is GWGR, as distinct from GPGR, which is prevalent in the United States and Europe (4). Subtype F is the most prevalent non-B variant and represents approximately 18% of HIV-1 isolates from southeastern Brazilian cities. An increasing prevalence of subtype C infections has also been recognized in Brazil's southern-most states (5–7).

Manaus, with 1.2 million inhabitants, is the largest city in the Amazon Basin. The country's first imported AIDS case was confirmed in this city in 1986 (8). Approximately 100,000 Indians, belonging to 10 distinct ethnic groups, live in the rain forests of this region. Manaus is the major port of entry and a trade center of the region, and its population has both commercial and family ties to the indigenous people. The HIV-1 serosurveys conducted to date among the Amazonian Indians have not indicated a rise in infections among this population (9).

In this paper, we describe the first genetic analysis of HIV-1 isolates collected from people living in the Brazilian rain forest. The primary goal of this research was to determine whether complex genetic patterns of HIV isolates that have been identified in the urban southeast are also being established in the interior Amazon Basin. Our results show that the epidemic in this region is composed of prototypic subtypes B and F, as well as multiple genomic mosaics combining genetic elements from both subtypes B and F. This genetic phenomenon was found in strains circulating among all risk groups, associated with patients at every clinical stage of disease.

MATERIALS AND METHODS

Sample Collection

From November 1996 through March 1997, approximately 2500 people visited the Counseling and Testing AIDS Center of Instituto de Medicina Tropical de Manaus. Among them, 31 were found to be seropositive for HIV-1 by results of testing with commercially available enzyme immunoassays (EIAs) (Abbott Laboratories, Abbott Park, IL, U.S.A.; Sanofi Pasteur, Paris, France), and HIV-1 Western blot (Cambridge Biotech, Worcester, MA, U.S.A.).

Polymerase Chain Reaction and Sequencing

Total nucleic acid was extracted from 200 μl of the buffy coat using a QIAamp Blood kit (Qiagen Inc., Chatsworth, CA, U.S.A.), according to the manufacturer's protocol. Nucleic acid was eluted in 50-μl water and stored at −70°C. Three different HIV-1 genomic regions were targeted for PCR amplification: *gag* p24, the env C2V3 region of gp120, and *pol* protease (*pro*). Nested polymerase chain reaction (PCR) was performed, as previously described (5,7). Restriction fragment-length polymorphism (RFLP) of the protease gene was used to putatively subtype the protease gene using the Alu I and Bcl I digestion patterns, following the protocols previously described in Janini et al. (5).

PCR products were purified using a QIAamp PCR purification kit (Qiagen) according to manufacturer's instructions. Double-stranded PCR fragments were directly sequenced in both directions with fluorescent dye-labeled sequencing terminators, using an automated sequencer (ABI Model 370, Perkin Elmer-Cetus, Norwalk, CT, U.S.A.).

Phylogenetic Analysis

Sequence data were analyzed using Sequencer 3.0 (Gene Codes Corp., Ann Arbor, MI, U.S.A.). Alignments were generated using DNASIS version 2.1 software (Hitachi, West Irvine, CA, U.S.A.). Evolutionary distances were estimated, using the Kimura two-parameter method, and relationships were determined using neighbor-joining methods. The sequence of $SIV_{cpz\ gab}$ was used as the outgroup. Sequences described in this study have been assigned the following GenBank accession numbers: pro (AF076439-AF076448), gag p24 (AF076323-AF07632), and env C2V3 (AF076313-AF076322).

RESULTS

The genomic DNA samples from the 31 HIV-infected people were isolated, and the C2V3 envelope regions were amplified by PCR, sequenced, and analyzed. Nearly equal proportions of subtype B (16 of 31 [51.6%]) and subtype F (15 of 31 [48.4%]) sequences were identified by C2V3 sequence analysis. All isolates were simultaneously analyzed by RFLP targeting the *pro* gene. Comparison of *env* subtyping and pro RFLP patterns suggested the presence of four distinct genomic structures among these samples. Of the 31 specimens, 12 (38.70%) exhibited consistent phylogenetic groupings across their genomes and were classified as pure prototypic subtype B; 5 (16.13 %) samples were similarly classified as pure subtype F; and 11 (35.40%) were characterized by potential mosaic genomes in which the pro gene and env C2V3 region clustered with different subtypes. The clinical and epidemiological characteristics of the individuals who participated in this study, as well as the genetic analysis of the two different genomic regions for each patient s sample are shown in Table 1. All genetic variants were present in different risk groups, and no association was found with clinical stages (World Health Organization classification system), CD4 counts, or viral loads. Of note, although all 14 mosaic genomes were identified among the 24 male participants, the 7 female participants were found to carry only the prototypic strains.

To clarify the actual genomic structure of these viruses, we selected 10 samples representing four putative genomic structures and analyzed by sequencing the *pro* and *gag* p24 regions. The phylogenetic trees representing

TABLE 1. *Summary of epidemiologic, clinical, and viral genotype data*

Code	Stage[a]	Gender	Risk factor[b]	Age (y)	Subtyping pro-env[c]	CD4 counts (cells/mm^3)	Viral load (RNA copies/ml)
BRAM05	B	M	HET	27	FB	105	NA
BRAM07	A	F	HET	19	FF	NA	NA
BRAM08	A	M	BI	42	BF	165	NA
BRAM09	C	F	HET	38	FF	248	NA
BRAM10	B	M	HOM	27	FB	99	NA
BRAM11	A	M	HET	45	BF	380	36,000
BRAM12	A	M	HET	40	BF	552	810
BRAM13	B	M	HET	25	FB	NA	NA
BRAM14	B	M	BI	26	BF	607	48,000
BRAM15	B	M	HET	33	BF	NA	NA
BRAM16	A	F	HET	31	BB	340	NA
BRAM17	B	M	BI	29	FF	976	<400
BRAM18	A	M	HOM	27	BF	679	14,000
BRAM19	C	M	HOM	30	BB	220	NA
BRAM20	C	M	BI	29	FF	477	NA
BRAM23	C	M	HOM	24	BF	NA	NA
BRAM24	B	M	BI	35	BB	1464	85,000
BRAM25	A	M	HET	22	BF	974	130,000
BRAM26	A	M	HOM	35	BF	206	NA
BRAM27	A	M	HET	24	BF	527	4400
BRAM28	A	M	HOM	27	BF	287	25,000
BRAM29	A	M	NA	24	FB	482	23,000
BRAM30	A	F	HET	25	BB	380	23,000
BRAM31	B	M	HET	38	BB	360	NA
BRAM32	B	M	HET	26	BB	239	NA
BRAM33	B	M	HOM	32	BB	NA	NA
BRAM34	A	F	HET	25	BB	498	85,000
BRAM35	C	M	HET	35	BB	146	>400
BRAM36	B	M	HOM	28	BB	488	13,000
BRAM37	C	F	TX	39	BB	237	NA
BRAM38	A	F	HET	49	BB	183	26,000

[a] Stage of disease based on WHO classification.
[b] The risk factor is described as follow: HET (heterosexual), HOM (homosexual), BI (bisexual), and TX (transfusion).
[c] Subtyping call of protease and envelope regions (see Material and Methods).
NA, data not available.

the *gag* p24, *pro*, and C2V3 *env* regions of these 10 samples are shown in Figure 1. The sequence analysis of three different genetic regions confirmed the existence of different chimeric genomes and identified four patterns of mosaic structures. The samples BRAM 8, BRAM 14, BRAM 27, and BRAM 28 were classified as B *gag* p24 /B *pro* /F C2V3*env*. Sample BRAM 29 was classified as F *gag* p24/F *pro* /B C2V3*env*. A third mosaic structure (B *gag* p24/F *pro* /B^{C2V3}*env*) was found in sample BRAM 5, and a fourth (B *gag* p24/C *pro*/B^{C2V3}*env*) in sample BRAM 13. The pure prototypic B strains (BRAM 19 and BRAM 24) and a pure prototypic F strain (BRAM 9) were also confirmed by the phylogenetic analysis of all three distinct genetic regions showing a consistent subtyping across their genomes.

DISCUSSION

In this study, we evaluated the subtype frequencies among 31 HIV-infected persons from Manaus, the largest community within the central rain forest of Brazil's Amazon Basin. Based only on envelope subtyping, we found equal representation of subtypes B and F, which is different from those reported from results of previous studies, conducted among AIDS patients residing in Brazil's southeastern cities, which showing proportions of subtype F strains ranging from 8%-18% (4–7). A higher proportion of subtype F viruses in Manaus, could best be explained by a founder effect, that is, a simultaneous introduction in this region of subtypes B and F and its recombinants in relatively equal proportions. This hypothesis is supported by similar range of genetic diversity in the C2V3 *env* region among the B and F isolates from Manaus samples (19.6% for B and 19.2% for F), contrasting with a relative high diversity of subtype B viruses compared with subtype F among isolates from the southeastern region of Brazil. In that area, the genetic divergence of subtype F is roughly half that of its subtype B counterparts (7). There is a lack of difference in subtype distribution between different risk group and

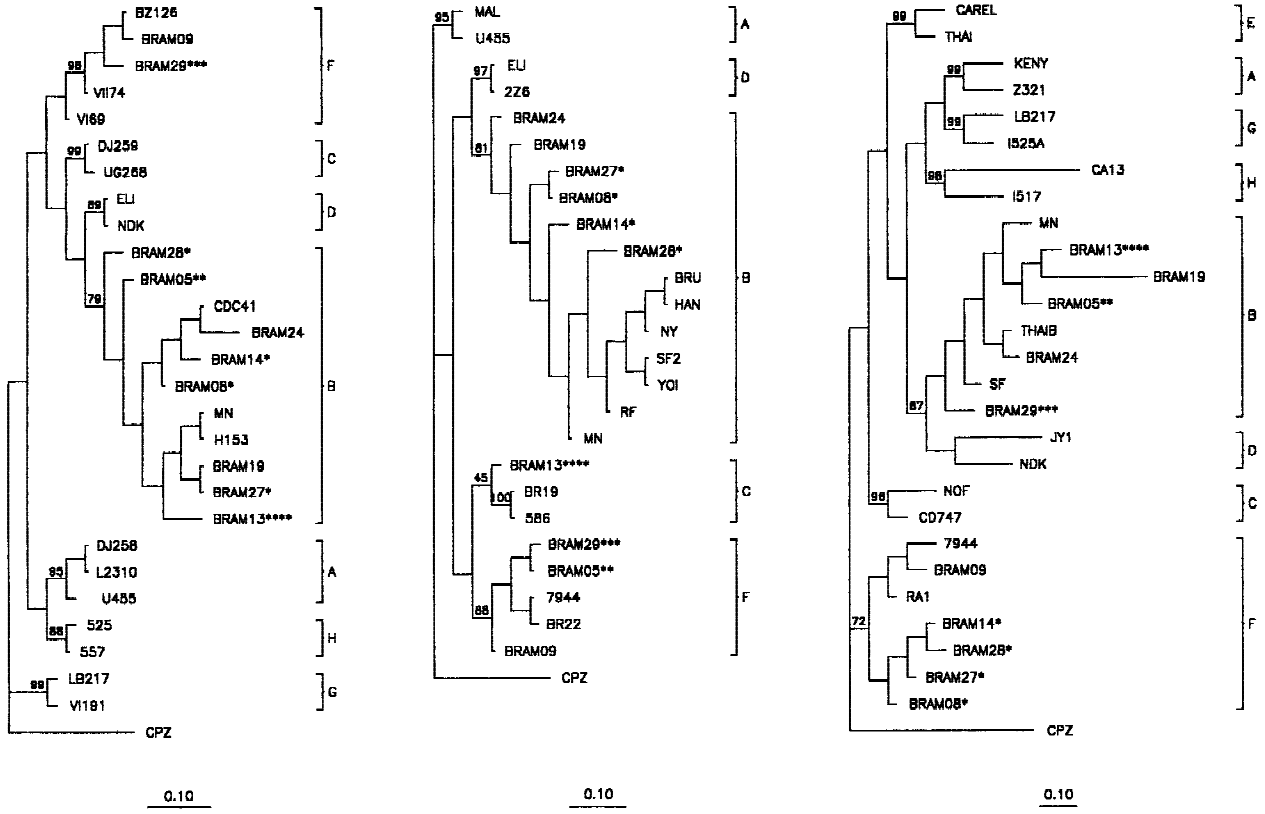

FIG. 1. Phylogenetic analysis of 10 representative Manaus HIV isolates compared with 20 reference HIV-1 group M subtypes available in the Los Alamos database using the neighbor-joining method (Kimura 2 parameters). Phenograms are shown in three panels: (**A**) *gag* p24; (**B**) protease gene (*pro*); and (**C**) *env* C2V3 aligned fragments. Sequence $SIV_{cpz\ gab}$ was used as an outgroup. Aligned fragments were analyzed as described in Materials and Methods, and bootstrap values for 100 replicates are listed at the major subtype branches. Scale bar indicates evolutionary distance of 0.10 nucleotides per position in sequences. Specimens of pure and mosaic genomes of Manaus isolates are designated as follows: no asterisks, pure B or F genomes; *, mosaic represented by subtype B p24-*gag* and *pro*, and subtype F *env* C2V3; **, mosaic represented by subtype B p24-*gag* and *env*C2V3, and subtype F *pro*; ***, mosaic represented by subtype F p24-*gag* and *pro*, and subtype B *env* C2V3; ****, mosaic represented by subtype B p24-*gag* and *env*C2V3, and subtype C *pro*.

clinical stage compared with situation in Thailand where the subtype E is clearly associated to injecting drug use (IDU) (10). The difference in subtype distribution observed between gender could be explained by the recent introduction of subtype F in this population, at the same time the epidemic was spreading through heterosexual contact, which produced high proportion of infected females with this viral variant. A similar finding was previously observed in Rio de Janeiro (11). However, because of the small number of people analyzed in this study ($n = 31$), it is to draw statistically significant conclusions.

The epidemic in Brazil serves as the largest single reservoir of non-B variants and recombinants in North and South America (12). Recombinant B/F genomes were recognized among the first isolates from Brazil (13). In one study, Cornelissen et al. studying several isolates from the WHO repository described both B/F and B/C recombinants in Brazil (14). Recombinant strains derived from B and F parental sequences have also been reported in Argentina (15). The cocirculation of multiple HIV-1 subtypes within a geographic region and the occurrence of dual infection of individuals with distinct viral subtypes (5,16) increase the probability of HIV-1 recombination the development of mosaic genomes, as has been observed in Brazil. The sequencing methods limited only to the analysis of a short viral genomic fragment (C2V3*env*) do not allow detect the occurrence of mosaic genomes of HIV. This study has demonstrated that complementary analysis of other genomic regions such as protease (*pro*) through RFLP improves the sensitivity to detect potential mosaic genomes. In this study, we found pure parental forms (subtype B and F), as well four mosaic patterns with different combinations of subtypes B, and F in the *gag* p24, *pro*, and *env* regions. A recombinant form between subtype B and C was observed, for which a parental subtype C strain was not found. Although both B and F parental

strains are prevalent in southeastern Brazil, the recombinant frequency observed in the Amazon Basin is substantially higher than that reported in Brazil s urban southeastern areas. We also observed the GWGR sequence in the "crown" of the *env* V3 loop was frequently preceded by a methionine residue. This motif is characteristic of many Brazilian subtype B strains but rarely found among subtype B isolates from the United States and Europe (4,6). Taken together, these data strongly suggest that the Amazonian subtype B strains were imported from the large cities in southeastern Brazil where similar variants circulate. Although this paper presents more data on HIV-1 strains in the Amazon region than any previous report, several points must be considered in interpreting our results. It is clear that the prevalence of HIV-1 among persons attending the counseling and testing center is fairly low and patients attending TB clinics may not be generalizable to all HIV-infected persons. Nevertheless, this is a complete sampling of all infected people who were seen at the site and probably represents those who were either infected in the Amazon region as well as those who were infected elsewhere and came to the area. The high proportion of mosaics is a very interesting result and suggests that these mosaics originated in a setting where the prevalence of two or more HIV-1 subtypes and the incidence of infections were both sufficiently high. This suggests that the subtype F recombinants, like the subtype B strains, were probably imported from southeastern Brazil through multiple individuals introductions.

The genetic diversity and presence of mosaic genomes among these samples reflect the complexity of the Brazilian HIV epidemic as a whole. The high proportion of recombinant viruses strongly suggests that recombination is an ongoing event and the nascent recombinant strains may acquire advantages for a wide geographic spread and play an important role in the future directions for the Brazilian epidemic.

Acknowledgments: This work was sponsored by the DST-AIDS Program of the Brazilian Ministry of Health, CNPq, FAPERJ, and FIOCRUZ.

REFERENCES

1. Myers G, Korber B, Foley B, et al. *Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences.* Los Alamos, NM: Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, 1996.
2. Hu DJ, Dondero TJ, Rayfield MA, et al. The emerging genetic diversity of HIV: the importance of global surveillance for diagnostics, research, and prevention. *JAMA* 1996;275:210–16.
3. Simon F, Mauclere P, Roques P, et al. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med* 1998;4:1032–7.
4. Morgado MG, Sabino EC, Shpaer EG, et al. V3 region polymorphism in HIV-1 from Brazil: prevalence of subtype B strains divergent from North American/European prototype and detection of subtype F. *AIDS Res Hum Retroviruses* 1994;10:569–76.
5. Janini LM, Pieniazek D, Peralta JM, et al. Identification of single and dual infections with distinct subtypes of human immunodeficiency virus type 1 by using restriction fragment length polymorphism analysis. *Virus Genes* 1996;13:69–81.
6. Sabino EC, Diaz RS, Brigido LF, et al. Distribution of HIV-1 subtypes seen in an AIDS clinic in Sao Paulo City, Brazil. *AIDS* 1996;10:1579–84.
7. Tanuri A, Swanson P, Sushil D, et al. HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil. *JAIDS J Acquir Immune Defic Syndr* 1999;20:60–6.
8. Barbosa N, Garrido M, Castilho MC, et al. Características espaciais do HIV/AIDS no estado do Amazonas. *Rev Soc Brasil Med Trop* 1998;31:122.
9. Andrada-Serpa MJ, Dobbin JA, Gomes P, et al. Incidence of retroviruses in some Brazilian groups. *Immunol Lett* 1988;18:15–18.
10. Ou C-Y, Takebe Y, Weiniger, et al. Independent introduction of two major HIV-1 genotypes into distinct high risk population in Thailand. *Lancet* 1993; 341:1171–4.
11. Pinto ME, Tanuri A, Schechter M. Molecular and epidemiological evidence for the discontinuous introduction of subtypes B and F Into Rio de Janeiro, Brazil. *J Acquir Immune Defic Syndr Hum Retrovirol* 1998; 19:310–12.
12. Flores I, Pieniazek D, Moran N, et al. Presence of HIV-1 subtype F in single and dual infections in Puerto Rico: a potential sentinel site for monitoring novel genetic HIV variants in North America. *Emerg Infect Dis* 1999;5:481–3.
13. Sabino EC, Shpaer EG, Morgado MG, 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;68:6340–6.
14. Cornelissen M, Kampinga G, Zorgdrager F, et al. Human immunodeficiency virus type 1 subtypes defined by env show high frequency of recombinant gag genes. *J Virol* 1996;70:8209–12.
15. Marquina S, Leitner T, Rabinovich RD, et al. Coexistence of subtypes B, F and a B/F env recombinant of HIV type 1 in Buenos Aires, Argentina. *AIDS Res Hum Retroviruses* 1996;12:1651–4.
16. Andrew W, Artenstein TC, VanCott JR, et al. Dual Infection with human immunodeficiency virus type 1 of distinct envelope subtypes in humans. *J Infect Dis* 1996;171:805–10.