AIDS Research and Human Retroviruses

AIDS Research and Human Retroviruses: http://mc.manuscriptcentral.com/aidsresearch

High HIV-1 Genetic Diversity in Patients from Northern Brazil

Journal:	*AIDS Research and Human Retroviruses*
Manuscript ID	AID-2016-0044.R1
Manuscript Type:	Full Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Costa, Carolina; Fundação de Medicina Tropical Dr Heitor Vieira Dourado. Secretaria de Estado da Saúde. Governo do Estado do Amazonas.; Universidade do Estado do Amazonas Costa de Oliveira, Cíntia Mara; Fundação de Medicina Tropical Dr Heitor Vieira Dourado. Secretaria de Estado da Saúde. Governo do Estado do Amazonas. Chehuan de Melo, Yonne Francis; Fundação de Medicina Tropical Dr Heitor Vieira Dourado. Secretaria de Estado da Saúde. Governo do Estado do Amazonas. Delatorre, Edson; Oswaldo Cruz Institute, Bello, Gonzalo; Oswaldo Cruz Institute-FIOCRUZ, Immunology Couto-Fernandez, José; Laboratório de AIDS & Imunologia Molecular. Instituto Oswaldo Cruz, FIOCRUZ
Keyword:	HIV, Virus Evolution/Diversity, phylogenetics

Title: High HIV-1 Genetic Diversity in Patients from Northern Brazil

Running title: HIV-1 Genetic Diversity in Northern Brazil

Authors: Carolina Marinho da Costa[1,2], Cintia Mara Costa de Oliveira[1], Yonne Francis Chehuan de Melo[1], Edson Delatorre[3], Gonzalo Bello[3]*, and Jose Carlos Couto-Fernandez[3]

1. Laboratório de Genotipagem de HIV. Fundação de Medicina Tropical Dr Heitor Vieira Dourado. Secretaria de Estado da Saúde. Governo do Estado do Amazonas. Manaus. Brazil.

2. Universidade do Estado do Amazonas. Manaus. Brazil.

3. Laboratório de AIDS & Imunologia Molecular. Instituto Oswaldo Cruz, FIOCRUZ. Rio de Janeiro. Brazil.

Keywords: HIV, subtypes, Northern region, Brazil.

***Corresponding author:**

Lab. de AIDS & Imunologia Molecular, Instituto Oswaldo Cruz - FIOCRUZ. Av. Brasil 4365, 21045-900 Rio de Janeiro, RJ, Brazil. Phone: +55 21 3865 8154. Fax: +55 21 3865 8173. E-mail:

ABSTRACT

The HIV-1 epidemic in Brazil is driven by subtypes B, F1, C and recombinants forms among those subtypes. The distribution of HIV-1 subtypes, however, may vary across different Brazilian regions and the molecular epidemiologic profile in Northern Brazil remains poorly explored. HIV-1 *pol* sequences were obtained from 305 patients failing antiretroviral therapy followed at outpatient clinics from five Northern Brazilian states. The most prevalent HIV-1 clade observed in the Northern Brazilian region was subtype B (81%), followed by BF1 recombinants (10%), subtype F1 (4%), subtype C (3%), BC recombinants (2%), and BU recombinants (1%). Although HIV-1 subtype B was the predominant HIV-1 clade in Northern Brazil, its prevalence greatly varies among different states, ranging from 63% in Rondônia to 92% in Acre. Among the 37 HIV-1 recombinants sequences detected in the Northern Brazilian region, nine (24%) displayed a URF structure, five (14%) a CRF28/29_BF-like structure, and four (11%) a CRF31_BC-like structure. Two other BF1 recombinant patterns were identified in 16 (43%) and three (8%) samples that may correspond to two potentially new CRFs_BF characteristic of the Northern region. This study reveals that despite the low spatial connectivity with other Brazilian regions, the genetic complexity of the HIV-1 epidemic in Northern Brazil is very high and that the molecular epidemiologic pattern may vary across different northern states, reflecting a complex epidemic with multiple independent viral introductions into this Brazilian region.

INTRODUCTION

Since the beginning of the AIDS epidemic in Brazil, until June 2015, about 800,000 AIDS cases were registered in the country.[1] The Southeastern region concentrates most (53.8%) of the total AIDS cases identified in Brazil, followed by the Southern (20.0%), Northeastern (14.6%), Central-Western (5.9%), and Northern (5.7%) regions. Epidemiological data also support a variable trend in the AIDS incidence rate over the last decade across Brazilian regions.[1] Whereas the incidence of newly reported AIDS cases showed a trend to decrease or stabilization in the Southeastern, Southern and Central-Western regions, it continues to increase in the Northern and Northeastern regions.

The Northern Brazilian region comprises the states of Acre (AC), Amapá (AP), Amazonas (AM), Pará (PA), Rondônia (RO), Roraima (RR) and Tocantins (TO). It is the largest Brazilian region, corresponding to 45% of the national territory, bordered by Brazilian Central-Western and Northeastern regions to the south and east, Bolivia and Peru to the west, and Colombia, Venezuela, Guyana, Suriname and French Guyana to the north. Mostly covered by the Amazon rainforest, it is also the most isolated and least densely populated Brazilian region, with about 15 million inhabitants (3.8 inhabitants per km^2). About 45.000 AIDS cases were reported in the Northern Brazilian region since detection of the first AIDS case in 1986 until June 2015.[1] The incidence of newly reported AIDS cases experienced a sharp increase in the Northern region over last years, rising from 14.3 cases per 100.000 inhabitants in 2005 to 25.7 cases per 100.000 inhabitants in 2014, well above the national mean (20.5 cases per 100.000 inhabitants).[1] Among the 10 Brazilian state's capitals with the highest incidence of newly reported AIDS cases in 2014, four were located in the Northern region: Manaus (AM), Porto Velho (RO), Belem (PA), and Boa Vista (RR), supporting the dynamic nature of the AIDS epidemic in this Brazilian region.

The Brazilian AIDS epidemic is mostly driven by HIV-1 subtypes B, F1, C and recombinants forms among those subtypes, although the relative prevalence of different HIV-1 subtypes and recombinants could greatly vary across different Brazilian regions.[2-5] Previous studies support a high prevalence of subtype B (≥75%) in Northern Brazil, followed by subtype F1 (2-14%), BF1 recombinants (3-10%), and subtype C (0-6%).[6-13] These estimates, however, should be interpreted with caution because no molecular data is available from some Northern Brazilian states (AC, RO and RR) and a relative low number of HIV-1 sequences have been subtyped from most others states (n < 100), with the only exception of PA (n = 558).

The objective of this study was to obtain a clearer picture of the HIV-1 molecular diversity in different states from Northern Brazil. To this end, we subtyped a total of 305 HIV-1 *pol* gene sequences sampled at five different states from the North region (AC, AM, PA, RO and RR).

MATERIALS AND METHODS

Study patients. A total of 305 HIV-1 *pol* sequences were recovered from patients failing antiretroviral therapy followed at outpatient clinics from the Public Health System distributed throughout five states (AC, AM, PA, RO and RR) of the Northern Brazilian region. Whole-blood samples from HIV-1 infected patients were sent for genotyping analysis at the Fundação de Medicina Tropical Dr Heitor Vieira Dourado (Manaus) and the Fundação Oswaldo Cruz (Rio de Janeiro), between 2010 and 2011, in the context of the Brazilian Network for HIV Genotyping (RENAGENO). A fragment of around 1,250pb encompassing the whole protease (PR) and part of the reverse transcriptase (RT) regions of *pol* gene was amplified and sequenced by using the TruGene HIV Genotyping kit

(Siemens, Munich, Germany), under conditions recommended by the manufacturers. The study was approved by the Ethics Committee of "Fundação de Medicina Tropical Dr Heitor Vieira Dourado" (Number CAAE: 09121512.8.0000.0005).

HIV-1 subtyping. HIV-1 subtypes were initially determined with the REGA HIV-1 Subtyping Tool 3.0 software[14] and later confirmed by phylogenetic and bootscanning analyses with HIV-1 reference sequences of subtypes A–D, F–H, J and K and of Brazilian Circulating Recombinant Forms (CRFs), retrieved from the Los Alamos HIV database (ww.hiv.lanl.gov). Maximum-likelihood (ML) phylogenetic trees were reconstructed with the PhyML 3.0 program[15] to verify the clustering of Brazilian HIV-1 sequences with subtypes and CRFs. ML trees were constructed under the GTR+I+G nucleotide substitution model, selected by jModeltest program,[16] and using the SPR branch-swapping algorithm for heuristic tree search and the approximate likelihood-ratio test (*a*LRT) to estimate the reliability of the obtained tree topology. All sequences were also subjected to bootscanning analyses using the Simplot 3.5.1 software[17] to identify possible recombination breakpoints. Bootstrap values supporting branching with HIV-1 reference sequences were determined in NJ trees constructed using the K2-parameter substitution model, based on 100 resamplings, with a 250 nt sliding window moving in steps of 10 bases. Initially, only HIV-1 reference sequences of "pure" subtypes were used for bootscanning analysis. If the recombination pattern found was similar to a particular CRF, then the bootscanning analysis was repeated adding the reference sequences of the putative CRF and removing the sequences of the parental subtypes. Finally, each fragment assigned to a particular subtype was subject to ML phylogenetic analysis to further confirm the genetic structure of recombinant viruses. Sequence fragments that did not cluster with any known HIV-1 group M subtype or CRF with *a*LRT support of ≥ 0.70 were defined as U.

RESULTS

Three hundred and five HIV-1 *pol* sequences from patients living in five different states from Northern Brazil (AC, AM, PA, RO and RR) were analyzed in the present study. This sample represents about 1% of the total number of people diagnosed with HIV/AIDS in those Brazilian states in the last 10 years.[1] Most patients were from AM (34%) and PA (33%), followed by RO (17%), RR (12%) and AC (4%). That distribution roughly matches the overall contribution of each state to the total number of AIDS cases diagnosed in the Northern Brazilian region: AM (31%), PA (43%), RO (10%), RR (4%) and AC (2%).[1] The most prevalent HIV-1 clade in our sample was subtype B (*n* = 248, 81%), followed by BF1 recombinants (*n* = 30, 10%), subtype F1 (*n* = 12, 4%), subtype C (*n* = 8, 3%), BC recombinants (*n* = 5, 2%), and BU recombinants (*n* = 2, 1%) (Figs. 1 and S1). According to the HIV-1 subtype distribution at *pol* (PR/RT) gene estimated here and in previous studies,[8, 13] the prevalence of most genetic variants varied over a large range across different Northern Brazilian states: subtype B (63-92%), subtype F1 (0-14%), subtype C (0-6%), BF1 (4-18%), and BC (0-8%) (Fig. 1). The Northern Brazilian state with the highest prevalence of non-subtype B variants was RO (37%), followed by AP (26%), AM (19%), TO (20%), RR (14%), PA (12%) and AC (8%).

A closer inspection of recombinant sequences revealed that the 30 BF1 recombinants detected were distributed in nine independent lineages (Figs. 2 and S2). One BF1 lineage comprises sequences sampled at AM, PA and RR (*n* = 5) that displayed the same mosaic structure at *pol* gene (Fig. 2) and branched with high support (Fig. S2) with the Circulating Recombinant Forms (CRFs) 28_BF and 29_BF reference sequences, and were thus classified as CRF28/29_BF-like recombinants. Two other BF1 lineages comprise more than two sequences with the same mosaic structure: BF-I (*n* = 16), and BF-II (*n* = 3) (Fig. 2).

The lineages BF-I (that comprises sequences from AC, AM, RO and RR) and BF-II (that only comprises sequences from RO) do not branch with any CRF_BF previously described in Brazil (Fig. S2) or in South America (data not shown), and probably represent new CRFs_BF characteristic of the Northern Brazilian region. The remaining BF1 lineages detected in Northern Brazil comprise only one sequence each and were classified as Unique Recombinant Forms (URFs_BF). All BC recombinants detected in RO displayed the same mosaic structure at *pol* gene (Fig. 2) and branched with high support (Fig. S2) with the CRF31_BC reference strains and were classified as CRF31_BC-like recombinants, whereas the other BC strain from AM was classified as a URF_BC. Finally, the two BU recombinants identified displayed unique mosaic structures (Fig. 2).

DISCUSSION

The AIDS epidemic in the Northern Brazilian region continues to expand, but the genetic characteristics of the HIV-1 strains circulating in this vast and relatively isolated region is largely unknown. The results from this study among patients from five different Northern Brazilian states indicate that the genetic complexity of the HIV-1 epidemic in the North is very high.

Despite the low spatial connectivity and population density, the HIV-1 epidemic in Northern Brazil resulted from multiple independent viral introductions and is dominated by subtype B, followed by F1 and BF1 recombinants, comparable to that found in the Southeastern, Northeastern and Central-Western Brazilian regions.[2-5] HIV-1 subtype C and BC recombinants were also detected in the Northern region, consistent with previous studies.[7-9, 11, 13] Although the overall prevalence of HIV-1 subtype C in Northern Brazil was low (<5%), this viral strain was detected in all northern states with a relative high sampling

size (n > 40 sequences) (Fig. 1). This supports the notion that HIV-1 subtype C has experienced a slow (but steady) northward expansion from the epicenter in the southernmost states,[18] and now reaches nearly all Brazilian states.

Recombinant strains comprise 13% of all HIV-1 *pol* sequences from Northern Brazil here analyzed. Some of the BF1 and BC recombinants sequences detected in the Northern region displayed a CRF28/29_BF-like and a CRF31_BC-like recombinant pattern. Although the CRF28/29_BF and CRF31_BC have been mostly associated to the Brazilian states of Sao Paulo (Southeastern region)[19-21] and Rio Grande do Sul (Southern region)[22-26], respectively, BC and BF1 *pol* sequences with a CRF-like mosaic structure have been described in some states from the Central-Western,[27-29] Northeastern,[30, 31] and Northern[13] Brazilian regions, supporting a wider dissemination of these Brazilian CRFs than originally described. The highest prevalence of CRF28/29_BF-like sequences was detected in the states of AM and RR (3%) and of CRF31_BC-like recombinants in RO (8%).

In this work we also detected for the first time the circulation of two potentially new CRFs_BF, here called lineage I and lineage II. The BF1 lineage I was particularly remarkable because it seems to reach a relative high prevalence in northern states located on the west: RO (10%), RR (8%), AC (8%) and AM (7%), and may represents an epidemiologically relevant regional CRF_BF. The BF1 lineage II, by contrast, was only detected in RO at a relative high prevalence (6%). The detection of these putative CRFs supports that HIV-1 epidemic in Northern Brazil has been also shaped by both viral dissemination through local (state-specific) transmission networks and viral exchanges among individuals from different Northern Brazilian states. Subsequent full-length genome analyses should be performed to determine whether those recombinants represent new CRFs_BF.

Although HIV-1 subtype B was the predominant HIV-1 clade in Northern Brazil, its prevalence greatly varies among different states, ranging from 63% in RO to 92% in AC. The states of AP and TO displayed the highest prevalence of subtype F1 (14%) and subtype C (6%), respectively, as described previously.[8, 13] Notably, despite the low size of the HIV epidemic in RO with only 4,723 AIDS cases described up to June 2015,[1] this state displayed a quite complex molecular epidemiologic profile and the highest prevalence of BF1 (18%) and BC (8%) strains among northern states. It is also interesting to note that despite the high prevalence of BF1 and BC variants detected in RO, most recombinant viruses branched within only three clades (BF1 lineage I, BF1 lineage II and CRF31_BC-like) and only one sequence from this state was classified as a URF.

In summary, the results from this study support a complex HIV-1 molecular epidemiologic profile in the Northern Brazilian region characterized by the co-circulation of subtypes B, F1, C and diverse recombinant forms among those subtypes. HIV-1 subtype B was the predominant clade in Northern Brazil, but its prevalence greatly varies among different states. Our results also supports that the overall genetic complexity of the HIV-1 epidemic in the Northern region has been mostly shaped by: 1) multiple introductions of viral strains from other Brazilian regions (country-specific lineages), 2) viral exchanges among Northern Brazilian states (regional-specific lineages), and 3) local expansion of some viral strains (state-specific lineages). This study represents an important contribution to understand the molecular epidemiological characteristics of the HIV-1 epidemic in the Amazon Brazilian region.

ACKNOWLEDGMENTS

E.D. is funded by a fellowship from "Programa Nacional de Pós-Doutorado (CAPES-Brazil)".

SEQUENCE DATA

Sequences were deposited in GenBank under accession numbers KU762018-KU762322.

REFERENCES

1. Brazilian. Ministry of Health. *AIDS Epidemiological Bulletin [in Portuguese]. January-June 2015; Ano IV, n° 01. Available from:* http://www.aids.gov.br/sites/default/files/anexos/publicacao/2015/58534/boletim_aids_11_2015_web_pdf_19105.pdf. 2015.

2. Brindeiro RM, Diaz RS, Sabino EC, et al. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. *Aids.* May 2 2003;17(7):1063-1069.

3. Inocencio LA, Pereira AA, Sucupira MC, et al. Brazilian Network for HIV Drug Resistance Surveillance: a survey of individuals recently diagnosed with HIV. *J Int AIDS Soc.* 2009;12:20.

4. Sprinz E, Netto EM, Patelli M, et al. Primary antiretroviral drug resistance among HIV type 1-infected individuals in Brazil. *AIDS Res Hum Retroviruses.* Sep 2009;25(9):861-867.

5. Alencar CS, Sabino EC, Carvalho SM, et al. HIV genotypes and primary drug resistance among HIV-seropositive blood donors in Brazil: role of infected blood

donors as sentinel populations for molecular surveillance of HIV. *J Acquir Immune Defic Syndr.* Jul 1 2013;63(3):387-392.

6. Vicente AC, Otsuki K, Silva NB, et al. The HIV epidemic in the Amazon Basin is driven by prototypic and recombinant HIV-1 subtypes B and F. *J Acquir Immune Defic Syndr.* Apr 1 2000;23(4):327-331.

7. Machado LF, Ishak MO, Vallinoto AC, et al. Molecular epidemiology of HIV type 1 in northern Brazil: identification of subtypes C and D and the introduction of CRF02_AG in the Amazon region of Brazil. *AIDS Res Hum Retroviruses.* Oct 2009;25(10):961-966.

8. Carvalho BC, Cardoso LP, Damasceno S, Stefani MM. Moderate prevalence of transmitted drug resistance and interiorization of HIV type 1 subtype C in the inland North State of Tocantins, Brazil. *AIDS Res Hum Retroviruses.* Oct 2011;27(10):1081-1087.

9. Cunha LK, Kashima S, Amarante MF, et al. Distribution of human immunodeficiency virus type 1 subtypes in the State of Amazonas, Brazil, and subtype C identification. *Braz J Med Biol Res.* Feb 2012;45(2):104-112.

10. Macêdo O, Ferreira LM, Lopes CAF, de Sousa RCM, Araujo JRM, Vasconcelos PFC. Distribution of HIV-1 subtypes in patients with HAART therapeutic failure in the States of Pará and Amazonas, Brazil: 2002 to 2006. *Rev Pan-Amaz Saude.* 2012;3(2):11-16.

11. de Moraes Soares CM, Vergara TR, Brites C, et al. Prevalence of transmitted HIV-1 antiretroviral resistance among patients initiating antiretroviral therapy in Brazil: a surveillance study using dried blood spots. *J Int AIDS Soc.* 2014;17:19042.

12. Lopes CA, Soares MA, Falci DR, Sprinz E. The Evolving Genotypic Profile of HIV-1 Mutations Related to Antiretroviral Treatment in the North Region of Brazil. *Biomed Res Int.* 2015;2015:738528.

13. Dos Anjos Silva L, Divino F, da Silva Rego MO, et al. HIV-1 Genetic Diversity and Transmitted Drug Resistance in Antiretroviral Treatment-Naive Individuals from Amapa State, Northern Brazil. *AIDS Res Hum Retroviruses.* Dec 1 2015.

14. de Oliveira T, Deforche K, Cassol S, et al. An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics.* Oct 1 2005;21(19):3797-3800.

15. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* May 2010;59(3):307-321.

16. Posada D. jModelTest: phylogenetic model averaging. *Mol Biol Evol.* Jul 2008;25(7):1253-1256.

17. Ray S. Simplot v2.5.0. *Available from:* http://www.med.jhu.edu/deptmed/sray/download/.

18. Graf T, Pinto AR. The increasing prevalence of HIV-1 subtype C in Southern Brazil and its dispersion through the continent. *Virology.* Sep 20 2012.

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

20. de Sa-Filho DJ, Soares Mda S, Candido V, et al. HIV type 1 pol gene diversity and antiretroviral drug resistance mutations in Santos, Brazil. *AIDS Res Hum Retroviruses.* Mar 2008;24(3):347-353.

21. de Souza AC, de Oliveira CM, Rodrigues CL, Silva SA, Levi JE. Short communication: Molecular characterization of HIV type 1 BF pol recombinants from Sao Paulo, Brazil. *AIDS Res Hum Retroviruses.* Dec 2008;24(12):1521-1525.

22. Santos AF, Sousa TM, Soares EA, et al. Characterization of a new circulating recombinant form comprising HIV-1 subtypes C and B in southern Brazil. *Aids.* Oct 24 2006;20(16):2011-2019.

23. Santos AF, Schrago CG, Martinez AM, et al. Epidemiologic and evolutionary trends of HIV-1 CRF31_BC-related strains in southern Brazil. *J Acquir Immune Defic Syndr.* Jul 1 2007;45(3):328-333.

24. Brigido LF, Nunes CC, Oliveira CM, et al. HIV type 1 subtype C and CB Pol recombinants prevail at the cities with the highest AIDS prevalence rate in Brazil. *AIDS Res Hum Retroviruses.* Dec 2007;23(12):1579-1586.

25. Passaes CP, Bello G, Lorete RS, et al. Genetic characterization of HIV-1 BC recombinants and evolutionary history of the CRF31_BC in Southern Brazil. *Infect Genet Evol.* Jul 2009;9(4):474-482.

26. de Medeiros RM, Junqueira DM, Matte MC, Barcellos NT, Chies JA, Matos Almeida SE. Co-circulation HIV-1 subtypes B, C, and CRF31_BC in a drug-naive population from Southernmost Brazil: analysis of primary resistance mutations. *J Med Virol.* Oct 2011;83(10):1682-1688.

27. Ferreira AS, Cardoso LP, Stefani MM. Moderate prevalence of transmitted drug resistance and high HIV-1 genetic diversity in patients from Mato Grosso State, Central Western Brazil. *J Med Virol.* Aug 2011;83(8):1301-1307.

28. Alcantara KC, Reis MN, Cardoso LP, Bello G, Stefani MM. Increasing heterosexual transmission of HIV-1 subtype C in Inland Central western Brazil. *J Med Virol.* Nov 21 2012.

29. da Silveira AA, Cardoso LP, Francisco RB, de Araujo Stefani MM. HIV type 1 molecular epidemiology in pol and gp41 genes among naive patients from Mato Grosso do Sul State, central western Brazil. *AIDS Res Hum Retroviruses.* Mar 2012;28(3):304-307.

30. 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.* Mar 2009;81(3):391-399.

31. Santos LA, Monteiro-Cunha JP, Araujo AF, Brites C, Galvao-Castro B, Alcantara LC. Detection of distinct human immunodeficiency virus type 1 circulating recombinant forms in northeast Brazil. *J Med Virol.* Dec 2011;83(12):2066-2072.

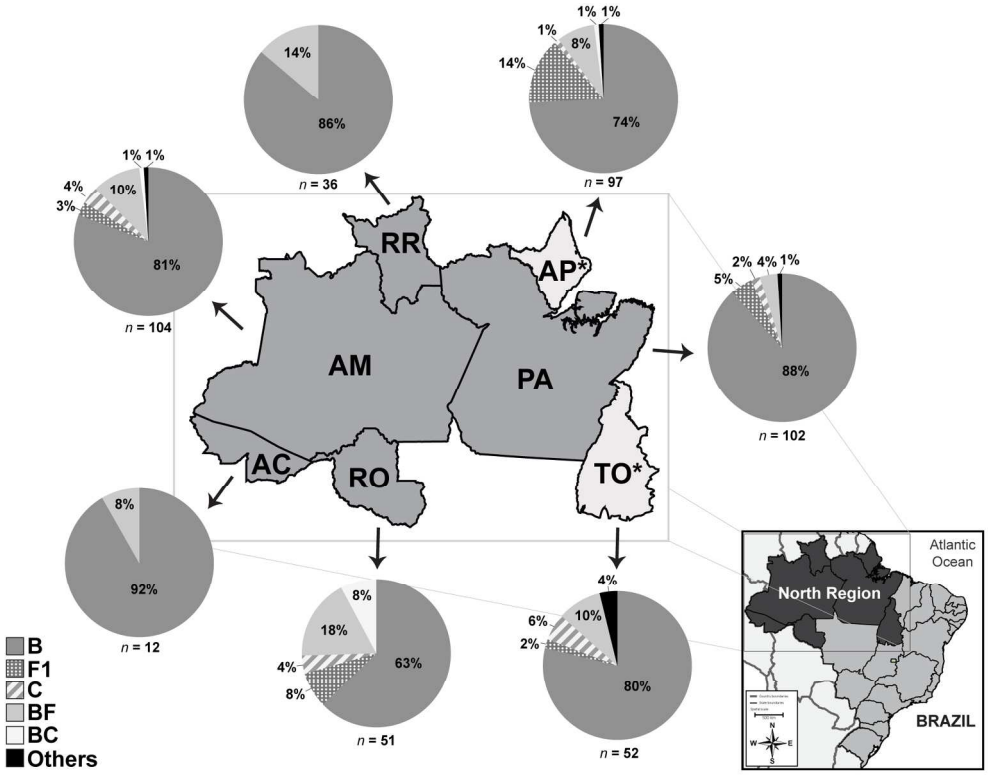


Figure 1. Map of Brazilian Northern region depicting the estimated prevalence of HIV-1 subtypes and inter-subtype recombinants among HIV-infected individuals from different states.
178x138mm (300 x 300 DPI)

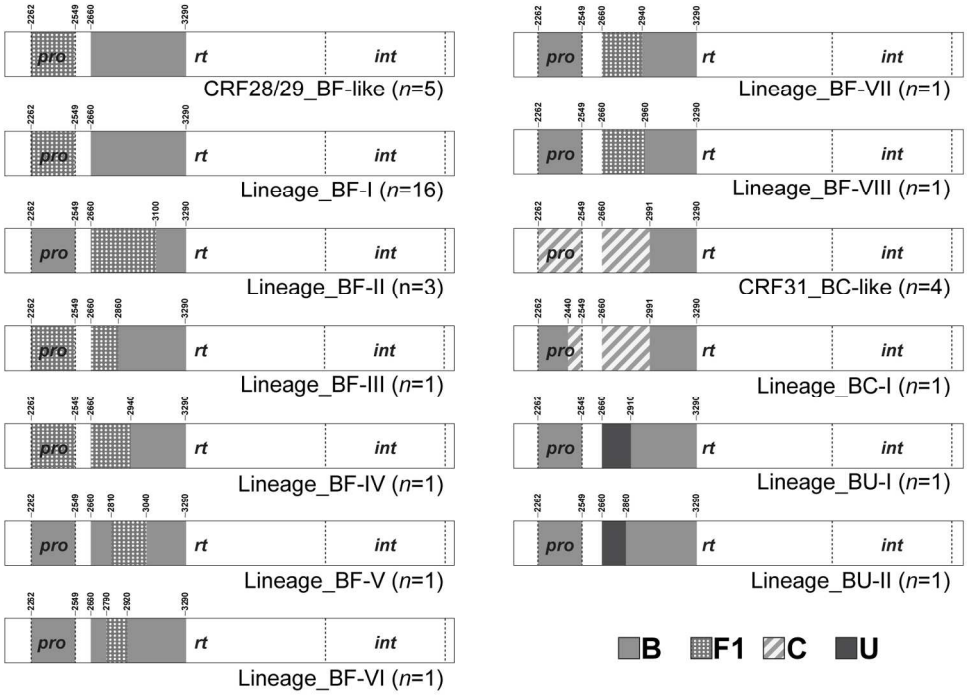
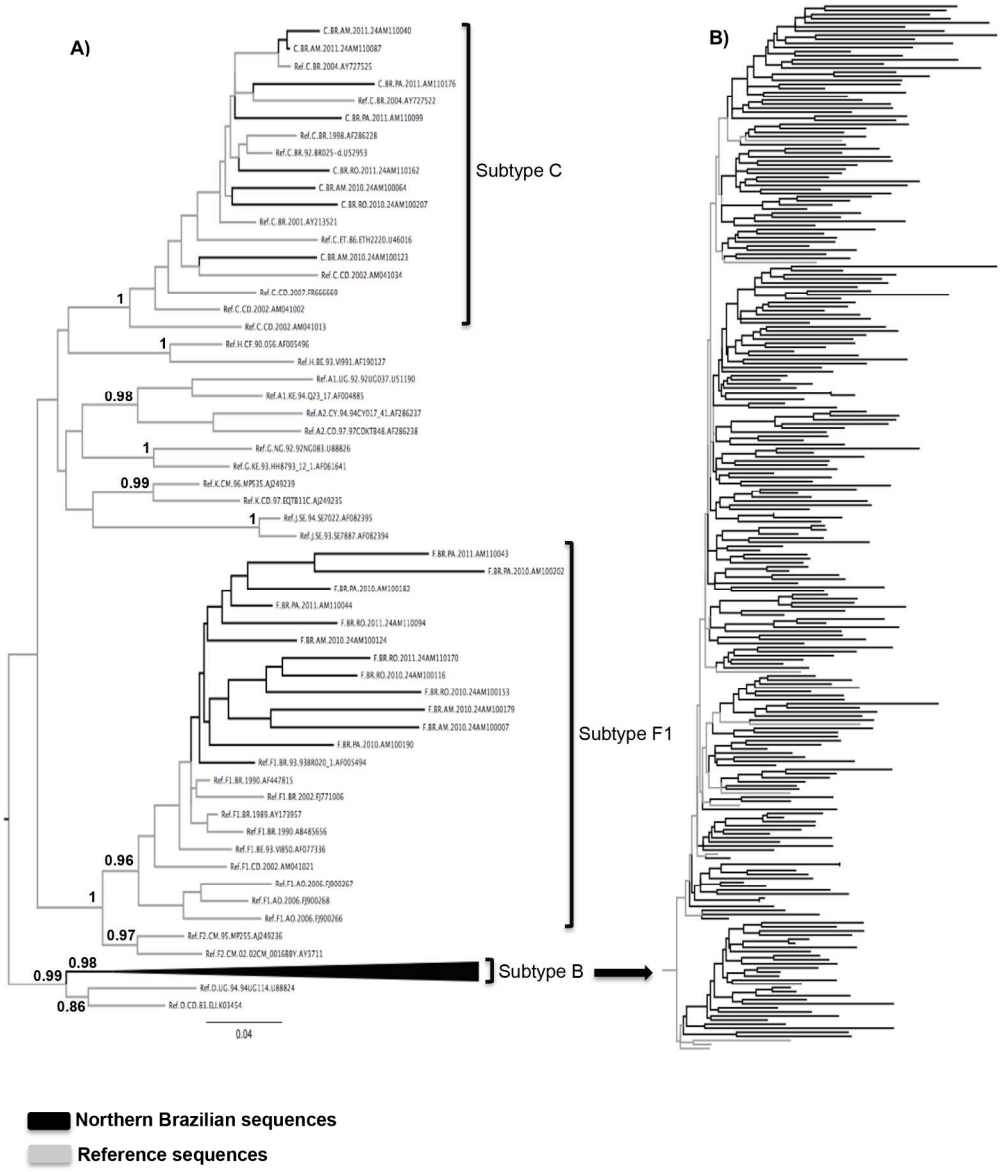
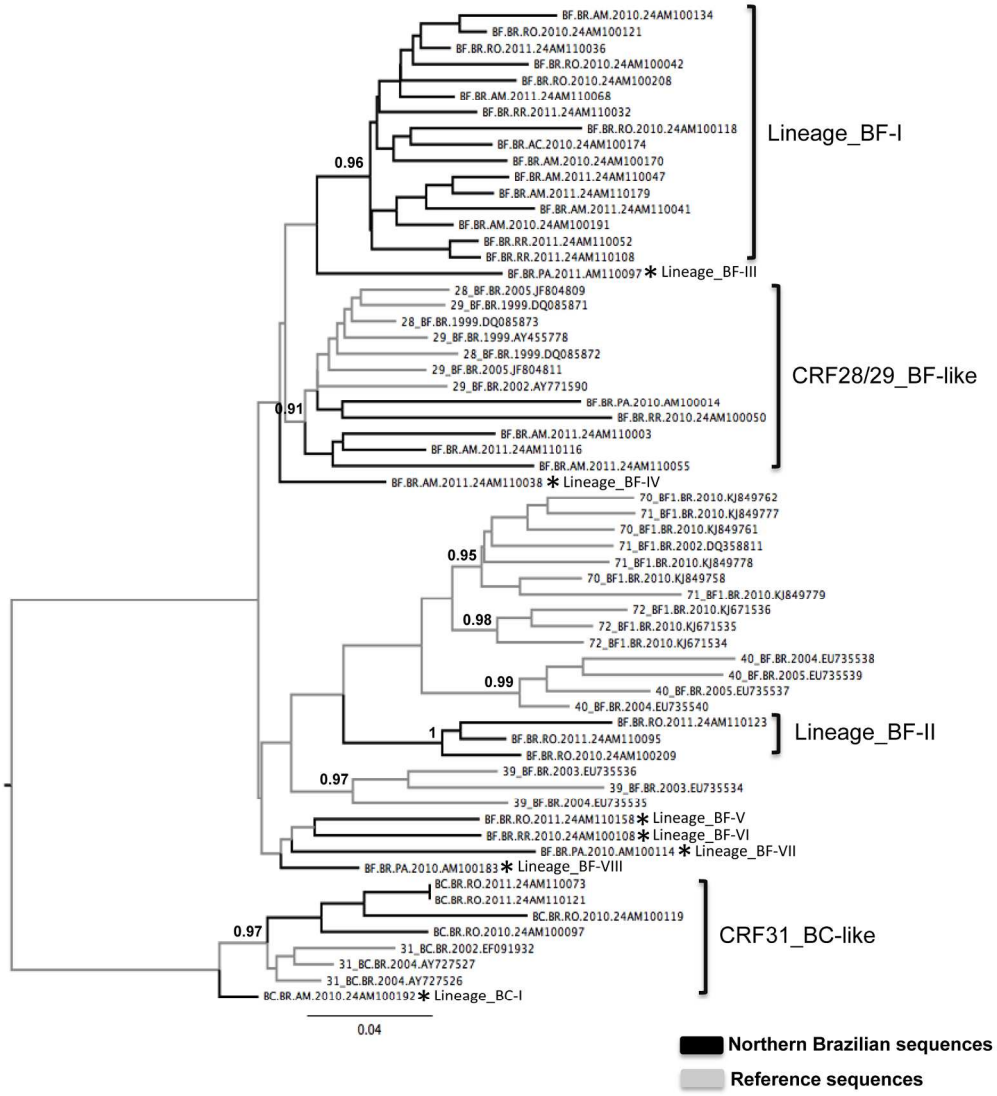


Figure 2. Schematic representation of the different mosaic patterns of inter-subtype recombinant HIV-1 pol sequences from Northern Brazil.
160x114mm (300 x 300 DPI)



190x225mm (300 x 300 DPI)



188x206mm (300 x 300 DPI)

Figure 1. Map of Brazilian Northern region depicting the estimated prevalence of HIV-1 subtypes and inter-subtype recombinants among HIV-infected individuals from different states. The total number of HIV-1 sequences analyzed in each locality is indicated bellow each graph. The HIV-1 subtype prevalence in the states of Amapá and Tocantins (marked with an asterisk) were estimated in previous studies [8,13]. AC: Acre. AM: Amazonas. AP: Amapá. PA: Pará. RO: Rondônia. RR: Roraima. TO: Tocantins.

Figure 2. Schematic representation of the different mosaic patterns of inter-subtype recombinant HIV-1 *pol* sequences from Northern Brazil. Segments are colored according to the subtype assignment as described in the legend. Fragments not analyzed were represented in white. Dotted lines delimitate the protease (*pro*), reverse transcriptase (*rt*) and integrase (*int*) regions.

Figure S1. A) Maximum likelihood phylogenetic tree of non-recombinant HIV-1 *pol* sequences from Northern Brazil. Sequences from Northern Brazil classified as "pure" subtypes (black branches) were combined with reference sequences of all HIV-1 group M subtypes (A-D, F-H, J and K) (grey branches). Brackets indicate the position of HIV-1 subtypes circulating in Brazil. The subtype B clade was collapsed for visual clarity. The aLRT branch support was shown only at key nodes corresponding to the most recent common ancestors of HIV-1 subtypes and sub-subtypes. Trees are rooted at midpoint and the branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site. B) Uncompressed subtree of subtype B clade showing in detail the distribution of Northern Brazilian sequences (black branches) among reference sequences (grey branches).

Figure S2. Maximum likelihood phylogenetic tree of inter-subtype recombinant HIV-1 *pol* sequences from Northern Brazil. Sequences from Northern Brazil classified as BC and BF1 recombinants (black branches) were combined with reference sequences of all Brazilian CRFs_BF (28, 29, 39 40, 70, 71 and 72) and CRF_BC (31) with recombination points at the *pol* gene fragment analyzed (grey branches). Brackets indicate the position of Brazilian CRF28/29_BF and CRF31_BC and of major local BF lineages detected in the Northern region. Asterisks point to HIV-1 sequences classified as URFs. The aLRT branch support was shown only at key nodes corresponding to the most recent common ancestors of HIV-1 CRFs and major

BF lineages. Trees are rooted at midpoint and the branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site.