

Identification of a Novel HIV-1 Circulating Recombinant Form (CRF72_BF1) in Deep Sequencing Data from Blood Donors in Southeastern Brazil

Rodrigo Pessôa,^a Anna Bárbara de Freitas Carneiro Proietti,^b Michael P. Busch,^c Sabri S. Sanabani^d

Department of Virology, São Paulo Institute of Tropical Medicine, University of São Paulo, São Paulo, Brazil^a; Foundation and Center for Hematology and Hemotherapy of Minas Gerais, Hemominas, Belo Horizonte, MG, Brazil^b; Blood Systems Research Institute, San Francisco, California, USA^c; Department of Pathology, Hospital das Clínicas, School of Medicine, University of São Paulo, São Paulo, Brazil^d

We report the identification of a novel HIV-1 circulating recombinant form (CRF72_BF1) in deep sequencing data from peripheral blood mononuclear cells (PBMC) of five blood donors in southeastern Brazil. Detection of this circulating recombinant form (CRF) confirms the need for effective surveillance to monitor the prevalence and distribution of HIV variants in a variety of settings in Brazil.

Received 8 April 2014 Accepted 28 May 2014 Published 12 June 2014

Citation Pessôa R, Carneiro Proietti ABDF, Busch MP, Sanabani SS. 2014. Identification of a novel HIV-1 circulating recombinant form (CRF72_BF1) in deep sequencing data from blood donors in southeastern Brazil. *Genome Announc*. 2(3):e00386-14. doi:10.1128/genomeA.00386-14.

Copyright © 2014 Pessôa et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Sabri S. Sanabani, sabyem_63@yahoo.com.

Both high mutation rates and recombination significantly contribute to the genetic diversification of human immunodeficiency virus type 1 (HIV-1). To date, HIV-1 viruses are classified into four phylogenetic groups, M, O, N, and P. The M group is further subdivided into nine subtypes (A to D, F to H, J, and K), among which subtypes A and F have been further classified into sub-subtypes (1, 2). Recombinant strains from at least three unlinked epidemiological sources, which exhibit identical mosaic patterns, have been classified separately as circulating recombinant forms (CRFs) (3). Recent estimates show that the HIV-1 CRFs and other minor recombinants account for approximately 20% of global HIV-1 infections (4). Here, we report the near full-length genome sequences (NFLGs) of a novel HIV-1 BF1 recombinant, designated CRF72_BF1 by the Los Alamos National Laboratory, derived from five blood donors in Minas Gerais, southeastern Brazil.

Cellular DNA was extracted from 5 peripheral blood mononuclear cells (PBMC) using the QIAamp blood kit (Qiagen) according to the manufacturer's instructions. The NFLGs from five overlapping fragments were obtained by PCR and determined by a previously reported method (5). The sequencing library was prepared as described previously (6). Briefly, the amplified fragments from a single viral genome were purified, quantified, and pooled at equimolar ratios. Approximately 1 ng of each pool was used in a fragmentation reaction. Finally, all libraries were pooled and loaded on an Illumina MiSeq for paired-end 250 sequencing. Fastq files were generated, validated, and *de novo* assembled into contiguous sequences and annotated with CLC Genomics Workbench version 5.5. The assembled contiguous sequences were aligned with reference sequences and screened for recombination by the bootscan methods (7, 8). Maximum likelihood trees were obtained by PhyML v. 3.1 using the GTR + I + G model (9). The approximate likelihood ratio test was used as a statistical test to calculate branch support.

The ultradeep sequencing yielded over 1.6×10^6 sequences reads, with average coverage ranging from $567 \times$ (10BR_MG025) to $14,333 \times$ (10BR_MG004). The NFLG consensus sequence from each strain of five BF1 (10BR_MG002, 10BR_MG003, 10BR_MG004, 10BR_MG008, and 10BR_MG025) was initially investigated using the bootscan method, which showed them to display identical mosaic structures, with 10 intersubtype breakpoints between subtype B and subclade F1 in *gag* (1 breakpoint), *pol* (4 breakpoints), *vif* (1 breakpoint), *env* (2 breakpoints), *rev2* (1 breakpoint), and *nef* (1 breakpoint), and with 6 subtype B regions and 5 subclade F1 regions. Comparison with published sequences revealed an additional two HIV sequences from Southeast Brazil (99UFRJ-1 and BREPM1029) that share an identical mosaic structure with those reported in our study (10, 11). Reconstructed trees for each region corroborated the results from the bootscan. The genetic distances between the CRF72_BF1 sequences exceed 8%, suggesting their longstanding circulation in Southeast Brazil.

The availability of the new CRF72_BF1 sequences described in this study should contribute to a more robust understanding of the overall genetic variability and phylogenetic relationships within and among other group M subtypes. Further molecular epidemiological investigations in a variety of settings are needed to identify the influence of CRF72_BF1 on the HIV epidemic in Brazil.

Nucleotide sequence accession numbers. All consensus genome assemblies generated in this study were submitted to NCBI's GenBank database under accession numbers [KJ671533](https://www.ncbi.nlm.nih.gov/nuccore/KJ671533) to [KJ671537](https://www.ncbi.nlm.nih.gov/nuccore/KJ671537).

ACKNOWLEDGMENTS

This work was supported by grants 2011/11090-5 and 2011/12297-2 from the Fundação de Amparo à Pesquisa do Estado de São Paulo.

We thank Brian T. Foley from the Los Alamos National Laboratory for advice and help on CRF nomenclature.

REFERENCES

1. Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S, Korber B. 2000. HIV-1 nomenclature proposal. *Science* 288:55–56. <http://dx.doi.org/10.1126/science.288.5463.55c>.
2. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Lemée V, Damond F, Robertson DL, Simon F. 2009. A new human immunodeficiency virus derived from gorillas. *Nat. Med.* 15:871–872. <http://dx.doi.org/10.1038/nm.2016>.
3. Carr JK, Salminen MO, Albert J, Sanders-Buell E, Gotte D, Birx DL, McCutchan FE. 1998. Full genome sequences of human immunodeficiency virus type 1 subtypes G and A/G intersubtype recombinants. *Virology* 247:22–31. <http://dx.doi.org/10.1006/viro.1998.9211>.
4. Hemelaar J, Gouws E, Ghys PD, Osmanov S, WHO-UNAIDS Network for HIV Isolation and Characterisation. 2011. Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS* 25:679–689. <http://dx.doi.org/10.1097/QAD.0b013e328342ff93>.
5. Sanabani S, Neto WK, de Sa Filho DJ, Diaz RS, Munerato P, Janini LM, Sabino EC. 2006. Full-length genome analysis of human immunodeficiency virus type 1 subtype C in Brazil. *AIDS Res. Hum. Retroviruses* 22:171–176. <http://dx.doi.org/10.1089/aid.2006.22.171>.
6. Pessoa R, Watanabe JT, Nukui Y, Pereira J, Kasseb J, Penalva de Oliveira AC, Segurado AC, Sanabani SS. 2014. Molecular characterization of human T-cell lymphotropic virus type 1 full and partial genomes by Illumina massively parallel sequencing technology. *PLoS One* 9:e93374. <http://dx.doi.org/10.1371/journal.pone.0093374>.
7. Salminen MO, Carr JK, Burke DS, McCutchan FE. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res. Hum. Retroviruses* 11:1423–1425. <http://dx.doi.org/10.1089/aid.1995.11.1423>.
8. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73:152–160.
9. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML. *Syst. Biol.* 3 0:59:307–321. <http://dx.doi.org/10.1093/sysbio/syq010>.
10. Thomson MM, Sierra M, Tanuri A, May S, Casado G, Manjon N, Najera R. 2004. Analysis of near full-length genome sequences of HIV type 1 BF intersubtype recombinant viruses from Brazil reveals their independent origins and their lack of relationship to CRF12_BF. *AIDS Res. Hum. Retroviruses* 20:1126–1133. <http://dx.doi.org/10.1089/aid.2004.20.1126>.
11. Sa-Filho D, Kallas EG, Sanabani S, Sabino E, Sucupira MC, Sanchez-Rosa AC, Tescarollo G, Tomiyama H, Bassichetto K, Janini LM, Diaz RS. 2007. Characterization of the full-length human immunodeficiency virus-1 genome from recently infected subjects in Brazil. *AIDS Res. Hum. Retroviruses* 23:1087–1094. <http://dx.doi.org/10.1089/aid.2006.0173>.