Genetic Characterization of HIV-1 Subtype D Near-Full-Length Proviral Genomes by Illumina Massively Parallel Sequencing Technology

Rodrigo Pessôa,[a] **Maria Esther Lopes,**[b] **Sabri S. Sanabani**[c]

Department of Virology, São Paulo Institute of Tropical Medicine, University of São Paulo, São Paulo, Brazil[a]; Hemorio, Rio de Janeiro, Brazil[b]; Department of Pathology, LIM 03, Hospital das Clínicas (HC), School of Medicine, University of São Paulo, São Paulo, Brazil[c]

This study describes the near-full-length genome deep sequencing of two HIV-1 subtype D strains identified in blood donors in Rio de Janeiro, Brazil, in what seems to have been a small restricted subtype D epidemic in the country.

Brazil, the world's fifth most populous country after China, accounts for about one-third of the HIV infections in Latin America, with an estimated 718,000 people infected (http://www.aids.gov.br). Most of these infections are caused by HIV-1 subtype B, except in the southern region, where subtype C prevails (1). HIV-1 subtype D viruses were first isolated from the peripheral blood lymphocytes in 1983 in patients from the Democratic Republic of the Congo (DRC) (2). Subtype D was also reported from Brazil in a dually infected individual in 1996 (3). Since then, there have been only sporadic cases with subtype D infection detected in the Rio de Janeiro (RJ) state (southeastern region) (4–7), but none were comprehensively sequenced. In this study, we report the first deep proviral genome sequencing of two HIV-1 subtype D variants obtained between 2007 and 2011 from Retrovirus Epidemiology Donor Study-II (REDS-II) blood donors in RJ.

Cellular DNA was extracted from 5 peripheral blood mononuclear cells (PBMC) using the QIAamp blood kit (Qiagen), according to the manufacturer's instructions. The near-full-length genomes (NFLGs) from five overlapping fragments were obtained by PCR and determined by a previously reported method (8). A sequencing library was prepared as described previously (9). Briefly, the amplified fragments from a single viral genome were purified, quantified, and pooled together at equimolar ratios. Approximately 1 ng of each pool was used in a fragmentation reaction. Finally, all libraries were pooled and loaded onto an Illumina MiSeq for paired-end 250-bp sequencing. Fastq files were generated, validated, and *de novo* assembled into contiguous sequences and annotated with CLC Genomics Workbench version 5.5. Maximum likelihood trees were obtained by PhyML version 3.1 using the GTR+I+G model (10). The approximate likelihood ratio test was used as a statistical test to calculate branch support.

The ultradeep sequencing yielded >1.6 × 10⁶ sequences reads, with average coverages ranging from 254× (10BR_RJ095) to 2,372× (10BR_RJ108). To determine the phylogenetic relationships of the newly characterized viruses, we constructed evolutionary trees from the NFLG consensus sequences. The results confirmed the initial diversity observed among subtype D previously described in the *pol* gene of HIV-1 (11). The intrasubtype distance for the two Brazilian variants was 8.1% and was comparable to the distances observed between subtypes D from different geographic locales. The close relationship of these Brazilian subtype D variants with sequences from Tanzania confirms an African origin for the subtype D circulation in Brazil. As inferred by geno2pheno coreceptor (12), both sequences were predicted to be X4 viruses.

This study describes the first NFLG HIV-1 subtype D viruses from South America. Despite early detection of subtype D in Brazil, it seemed not to have spread in much the same epidemic proportions as did subtype B or BF1 infections, which might imply that it was introduced and contained in only small networks.

Nucleotide sequence accession numbers. All consensus genome assemblies generated in this study were submitted to NCBI's GenBank database under accession no. KJ787683 and KJ787684.

REFERENCES

1. **Gräf T, Pinto AR.** 2013. The increasing prevalence of HIV-1 subtype C in southern Brazil and its dispersion through the continent. Virology **435:** 170–178. http://dx.doi.org/10.1016/j.virol.2012.08.048.
2. **Alizon M, Wain-Hobson S, Montagnier L, Sonigo P.** 1986. Genetic variability of the AIDS virus: nucleotide sequence analysis of two isolates from African patients. Cell **46:**63–74. http://dx.doi.org/10.1016/0092-8674(86)90860-3.
3. **Janini LM, Pieniazek D, Peralta JM, Schechter M, Tanuri A, Vicente AC, dela Torre N, Pieniazek NJ, Luo CC, Kalish ML, Schochetman G, Rayfield MA.** 1996. Identification of single and dual infections with distinct subtypes of human immunodeficiency virus type 1 by using restriction fragment length polymorphism analysis. Virus Genes **13:**69–81. http://dx.doi.org/10.1007/BF00576981.
4. **Morgado MG, Guimaraes ML, Gripp CB, Costa CI, Neves I, Jr, Veloso VG, Linhares-Carvalho MI, Castello-Branco LR, Bastos FI, Kuiken C, Castilho EA, Galvao-Castro B, Bongertz V.** 1998. Molecular epidemiology of HIV-1 in Brazil: high prevalence of HIV-1 subtype B and identifi-

cation of an HIV-1 subtype D infection in the city of Rio de Janeiro, Brazil. Evandro Chagas Hospital AIDS Clinical Research Group. J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. **18:**488–494.

5. **Guimarães ML, dos Santos Moreira A, Loureiro R, Galvão-Castro B, Morgado MG, Brazilian Network for HIV Isolation and Characterization.** 2002. High frequency of recombinant genomes in HIV type 1 samples from Brazilian southeastern and southern regions. AIDS Res. Hum. Retroviruses **18:**1261–1269. http://dx.doi.org/10.1089/088922202320886307.

6. **Brígido LF, Franco HM, Custódio RM, Oliveira CA, P Ferreira JL, Eira M, Bergel F, Araújo F, Carvalheiro JR, Rodrigues R.** 2005. Molecular characteristics of HIV type 1 circulating in São Paulo, Brazil. AIDS Res. Hum. Retroviruses **21:**673–682. http://dx.doi.org/10.1089/aid.2005.21.673.

7. **Couto-Fernandez JC, Eyer-Silva WA, Guimarães ML, Chequer-Fernandez SL, Grinsztejn B, Delaporte E, Peeters M, Morgado MG.** 2006. Phylogenetic analysis of Brazilian HIV type 1 subtype D strains: tracing the origin of this subtype in Brazil. AIDS Res. Hum. Retroviruses **22:**207–211. http://dx.doi.org/10.1089/aid.2006.22.207.

8. **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.

9. **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.

10. **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 3.0. Syst. Biol. **59:** 307–321. http://dx.doi.org/10.1093/sysbio/syq010.

11. **Alencar CS, Sabino EC, Carvalho SM, Leao SC, Carneiro-Proietti AB, Capuani L, Oliveira CL, Carrick D, Birch RJ, Gonçalez TT, Keating S, Swanson PA, Hackett J, Jr, Busch MP, NHLBI Retrovirus Epidemiology Donor Study-II (REDS-II), International Component.** 2013. 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. **63:**387–392. http://dx.doi.org/10.1097/QAI.0b013e31828ff979.

12. **Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R.** 2007. Bioinformatics prediction of HIV coreceptor usage. Nat. Biotechnol. **25:** 1407–1410. http://dx.doi.org/10.1038/nbt1371.