



Complete Endogenous Retrovirus Genome Sequence from a Brazilian Vampire Bat (*Desmodus rotundus*)

Luciano Chaves Franco Filho,^a Rafael Ribeiro Barata,^a Jedson Ferreira Cardoso,^a Janaina Mota de Vasconcelos Massafra,^c Poliana da Silva Lemos,^a Livia Medeiros Neves Casseb,^b Ana Cecilia Ribeiro Cruz,^b Marcio Roberto Teixeira Nunes^a

^aCentro de Inovações Tecnológica, Instituto Evandro Chagas, Ananindeua, Pará, Brazil

^bSeção de Arbovirologia e Febres Hemorrágicas, Instituto Evandro Chagas, Ananindeua, Pará, Brazil

^cInstituto de Ciências Biológicas, Universidade Federal do Pará (UFPA), Belém, Pará, Brazil

ABSTRACT The strain *Desmodus rotundus endogenous retrovirus* (DrERV) QR09 was obtained from a bat tissue sample collected from *Desmodus rotundus* in the Brazilian rain forest. The complete genome was sequenced using the next-generation sequencing strategy. The full-length genome of DrERV QR09 is 8,256 nucleotides in length and showed high similarity with other DrERVs.

Endogenous retroviruses (ERVs) are found in a wide variety of hosts and are possibly one of the first circulating viral strains of their respective hosts (1–3). ERVs that have integrated during host speciation, considered genetic fossils, are often used as markers for understanding the long-term phylogeny of the virus (4). In addition, several studies have shown that even when ERVs are inserted into the genome and considered inactive, there are indications that some proteins continue to be translated, even in low copies, and probably play important roles in cancer development and other gene regulatory activities (5–9).

The strain QR09 of *Desmodus rotundus endogenous retrovirus* (DrERV) was sequenced from a brain tissue sample of a bat (*Desmodus rotundus*) collected in the Brazilian Amazon rainforest (Viseu, Pará; 01°11'48"S, 46°08'24"W). The project was approved by the Ethics Committee on the Use of Animals of Evandro Chagas Institute (CEUA/IEC; number 031/2014) and Biodiversity Information and Authorization System (SISBIO; number 47592-1).

The viral particles were released from the cells using a stainless bead with a TissueLyser II (Qiagen), and the sample was pre-enriched using 0.45- μ m filters and treatment with benzonase (25 U/liter). DNA and RNA were extracted with a iPrep PureLink virus kit (Thermo Fisher) following the manufacturer's guidelines. The extracted DNA and RNA were quantified with a Qubit 2.0 fluorometer (Thermo Fisher) using the Qubit RNA HS assay kit, as well as the Qubit double-stranded DNA (dsDNA) HS assay kit (Thermo Fisher). The RNA samples were subjected to reverse transcription using the cDNA synthesis system kit (Roche, Branford, CT, USA), according to the manufacturer's guidelines. The cDNA and DNA of brain tissue were combined and sequenced as a single sample. Sequencing libraries were constructed using the Illumina Nextera XT DNA sample preparation kit and sequenced on an Illumina HiSeq 2500 instrument with the high-output V4 2 \times 100-bp sequencing kit (Table 1).

The raw data were filtered for Q30 quality, adapters were removed using Trim_galore pipeline v.0.4.5 (10), and reads less than 100 bp were removed using the Prinseq-lite.pl tool (11). The removal of the rRNA sequences was performed using the SortMeRNA tool v.2.1b (12). The assembly was performed by IDBA-UD v.1.1.3 (13) using default settings. The comparison with the protein nonredundant protein database was carried out by the DIAMOND tool v.0.9.22.123 (14), with an E value of 0.00001 (15). The results were annotated using the Blast2GO tool (16).

Citation Franco Filho LC, Barata RR, Cardoso JF, Massafra JMDV, Lemos PDS, Casseb LMN, Cruz ACR, Nunes MRT. 2019. Complete endogenous retrovirus genome sequence from a Brazilian vampire bat (*Desmodus rotundus*). *Microbiol Resour Announc* 8:e01497-18. <https://doi.org/10.1128/MRA.01497-18>.

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Copyright © 2019 Franco Filho et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Luciano Chaves Franco Filho, lucianofranco6@gmail.com.

Received 2 November 2018

Accepted 26 November 2018

Published 10 January 2019

TABLE 1 DrERV strain QR09 sequencing information

Parameter	Strain information
Bat species	<i>Desmodus rotundus</i>
Tissue	Brain
Sequencing platform	Illumina
DNA/cDNA input (ng/ μ l)	1
Total no. of reads	241,897,814
Total no. of contigs	407,850
N_{50} (bp)	19,212

The complete genome of DrERV strain QR09 is 8,256 nucleotides (nt) long, with 65-fold coverage and 48.6% GC content. Compared with the common structure of *Betaretrovirus*, the lineage shows the 4 common genes *gag* (2,187 nt), protease (837 nt), *pol* (2,418 nt), and *env* (1,770 nt). The *gag* and protease open reading frames did not present any stop codon; however, for the two coding regions *pol* and *env*, 2 and 3 stop codons were found, respectively. According to NCBI BLASTn analysis, the DrERV QR09 possesses 99% sequence identity over 100% query coverage of the DrERV isolate 824 (GenBank accession number [KP175580](#)) and 99% sequence identity over 96% query coverage of the DrERV isolate 216 (GenBank accession number [KP175581](#)); both genomes described were collected in Mexico from *D. rotundus*.

Data availability. The complete genome sequence of *Desmodus rotundus endogenous retrovirus* (DrERV) strain QR09 has been deposited in NCBI GenBank under the accession number [MH648003](#). The sequencing reads (under Sequence Read Archive number [SRR8208870](#)) can be accessed through BioProject number [PRJNA480298](#).

ACKNOWLEDGMENTS

This work was supported by Evandro Chagas Institute (IEC-PA Brazil) grants, MRTN CNPq (National Council for Scientific and Technological Development grant number 302584/2015-3), Fundação para o Desenvolvimento Científico e Tecnológico em Saúde (FIOTEC; grant number PRES-012-FIO-16), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; grant number 3274/2013), which provided a Ph.D. fellowship to Programa de Pós-Graduação em Virologia (PPGV-IEC), Ananindeua, Pará, Brazil.

REFERENCES

- Cui J, Tachedjian G, Tachedjian M, Holmes EC, Zhang S, Wang LF. 2012. Identification of diverse groups of endogenous gammaretroviruses in mega- and microbats. *J Gen Virol* 93:2037–2045. <https://doi.org/10.1099/vir.0.043760-0>.
- Cui J, Tachedjian M, Wang L, Tachedjian G, Wang LF, Zhang S. 2012. Discovery of retroviral homologs in bats: implications for the origin of mammalian gammaretroviruses. *J Virol* 86:4288–4293. <https://doi.org/10.1128/JVI.06624-11>.
- Zhuo X, Rho M, Feschotte C. 2013. Genome-wide characterization of endogenous retroviruses in the bat *Myotis lucifugus* reveals recent and diverse infections. *J Virol* 87:8493–8501. <https://doi.org/10.1128/JVI.00892-13>.
- Aswad A, Katzourakis A. 2016. Paleovirology: the study of endogenous viral elements, p 273–292. In Weaver SC, Denison M, Roossinck M, Vignuzzi M (ed), *Virus evolution*. Caister Academic Press, Poole, United Kingdom.
- Magiorkinis G, Katzourakis A, Lagiou P. 2017. Roles of endogenous retroviruses in early life events. *Trends Microbiol* 25:876–877. <https://doi.org/10.1016/j.tim.2017.09.002>.
- Katzourakis A, Aswad A. 2016. Evolution: endogenous viruses provide shortcuts in antiviral immunity. *Curr Biol* 26:427–429. <https://doi.org/10.1016/j.cub.2016.03.072>.
- Kassiotis G, Stoye JP. 2017. Making a virtue of necessity: the pleiotropic role of human endogenous retroviruses in cancer. *Philos Trans R Soc Lond B Biol Sci* 372:20160277. <https://doi.org/10.1098/rstb.2016.0277>.
- Frank JA, Feschotte C. 2017. Co-option of endogenous viral sequences for host cell function. *Curr Opin Virol* 25:81–89. <https://doi.org/10.1016/j.coviro.2017.07.021>.
- Blanco-Melo D, Gifford RJ, Bieniasz PD. 2017. Co-option of an endogenous retrovirus envelope for host defense in hominid ancestors. *Elife* 6:e22519. <https://doi.org/10.7554/eLife.22519>.
- Krueger F. Trim Galore. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
- Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28:3211–3217. <https://doi.org/10.1093/bioinformatics/bts611>.
- Peng Y, Leung HC, Yiu SM, Chin FY. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28:1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>.
- Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.
- Kerfeld AC, Scott KM. 2011. Using BLAST to Teach “E-value-tionary” concepts. *PLoS Biol* 9:e1001014. <https://doi.org/10.1371/journal.pbio.1001014>.
- Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36:3420–3435. <https://doi.org/10.1093/nar/gkn176>.