



# Complete Genome Sequence of Lleida Bat Lyssavirus

Denise A. Marston,<sup>a,c</sup> Richard J. Ellis,<sup>b</sup> Emma L. Wise,<sup>a</sup> Nidia Aréchiga-Ceballos,<sup>d,e</sup> Conrad M. Freuling,<sup>f</sup> Ashley C. Banyard,<sup>a</sup> Lorraine M. McElhinney,<sup>a,g</sup> Xavier de Lamballerie,<sup>c</sup> Thomas Müller,<sup>f</sup> Anthony R. Fooks,<sup>a,g,h</sup> Juan E. Echevarría<sup>d,i</sup>

Wildlife Zoonoses & Vector-Borne Diseases Research Group, Animal and Plant Health Agency (APHA), Surrey, United Kingdom<sup>a</sup>; Surveillance and Laboratory Services Department, APHA, Surrey, United Kingdom<sup>b</sup>; UMR\_D 190, Emergence de Pathologies Virales, Aix Marseille University, Marseille, France<sup>c</sup>; Instituto de Salud Carlos III, Madrid, Spain<sup>d</sup>; Laboratorio de Rabia, Instituto de Diagnóstico y Referencia Epidemiológicos, Ciudad de México, Mexico<sup>e</sup>; Friedrich Loeffler Institute (FLI), Institute of Molecular Virology and Cell Biology, Greifswald-Insel Riems, Germany<sup>f</sup>; Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom<sup>g</sup>; Institute for Infection and Immunity, St. George's Hospital Medical School, University of London, London, United Kingdom<sup>h</sup>; Ciber de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain<sup>i</sup>

**ABSTRACT** All lyssaviruses (family *Rhabdoviridae*) cause the disease rabies, an acute progressive encephalitis for which, once symptoms occur, there is no effective cure. Using next-generation sequencing, the full-genome sequence for a novel lyssavirus, Lleida bat lyssavirus (LLEBV), from the original brain of a common bent-winged bat has been confirmed.

The lyssavirus genome consists of a single-stranded negative-sense RNA of approximately 12 kb. The International Committee for the Taxonomy of Viruses recognizes 14 distinct lyssavirus species, with Lleida bat lyssavirus (LLEBV) and Gannoruwa bat lyssavirus (GBLV) awaiting classification (1, 2).

LLEBV was detected in a common bent-winged bat (*Miniopterus schreibersii*) found in Lleida, Spain, during July 2011 (1). Initial analysis of the partial nucleoprotein (N) sequence confirmed that this virus was most closely related to members of the *Lyssavirus* genus, grouping with West Caucasian bat lyssavirus and Ikoma lyssavirus sequences. Full-genome sequencing of the original brain material was undertaken to genetically characterize this virus. Generation of the full-genome sequence was particularly significant, as initial attempts at virus isolation failed. Viral RNA was extracted using TRIzol, and host genomic DNA and rRNA were depleted as described previously (3). Double-stranded cDNA (ds-cDNA) was synthesized using random hexamers and a cDNA synthesis system (Roche). The ds-cDNA was purified using AMPure XP magnetic beads (Beckman Coulter, Inc.), and 1 ng was used with the Nextera XT DNA sample preparation kit (Illumina), according to the manufacturer's instructions (omitting the bead normalization step) and sequenced using an Illumina MiSeq with 2 × 150-bp paired-end reads.

Mapping the sequencing data (6,651,896 reads) to the most closely related lyssavirus genome (Ikoma lyssavirus [IKOV]) failed to generate a consensus sequence. Therefore, the reads were processed to remove host genome by mapping to a mouse reference (sufficient homology to bat) using BWA version 0.7.5a-r405 (4). The remaining unmapped reads were assembled using IVA (5). Two virus-specific contiguous sequences (contigs) were obtained, which were confirmed by mapping the original data with BWA and consensus calling with a modified SAMtools script (6). The total number of assembled viral reads was 3,610 (0.25% of the host-depleted reads). The two contigs covered all coding regions, but a section of the G-L region was absent, and the genomic ends were atypical lengths. The G-L intergenic region and genomic ends were con-

Received 24 October 2016 Accepted 8 November 2016 Published 12 January 2017

**Citation** Marston DA, Ellis RJ, Wise EL, Aréchiga-Ceballos N, Freuling CM, Banyard AC, McElhinney LM, de Lamballerie X, Müller T, Fooks AR, Echevarría JE. 2017. Complete genome sequence of Lleida bat lyssavirus. *Genome Announc* 5:e01427-16. <https://doi.org/10.1128/genomeA.01427-16>.

© Crown copyright 2017. 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 Anthony R. Fooks, Tony.Fooks@apha.gsi.gov.uk.

firmed by Sanger sequencing of PCR amplicons, obtained using primers designed from the next-generation sequencing (NGS) consensus sequences, and paired with conserved primers that bind the conserved lyssavirus genomic termini (7). The sequence obtained by combining the two NGS contigs and three PCR amplicon sequences was used to map the original reads and determine the consensus sequence as described above. The total number of assembled viral reads for LLEBV was 12,629 (0.19% of total reads). The entire genome, apart from the last two nucleotides, was covered by NGS reads.

The genetic organization of the LLEBV genome is similar to that of other lyssaviruses, with a complete genome size of 11,931 nucleotides (nt). The open reading frame (ORF) lengths are as follows: 3' untranslated region (UTR), 70 nt; N-ORF, 1,353 nt; N-P intergenic region, 68 nt; P-ORF, 870 nt; P-M intergenic region, 74 nt; M-ORF, 609 nt; M-G intergenic region, 198 nt; G-ORF, 1,578 nt; G-L intergenic region, 608 nt; L-ORF, 6,381 nt; and 5' UTR, 122 nt. These data will contribute to our understanding of lyssavirus diversity and evolution and further our knowledge of vaccine-induced immunity and protection.

**Accession number(s).** The complete genomic sequence of LLEBV has been deposited in GenBank under accession number [KY006983](https://www.ncbi.nlm.nih.gov/nuccore/KY006983).

## ACKNOWLEDGMENTS

This project received funding from Defra, the Scottish Government, and the Welsh Government through grants SE0427 and SE0431 and from the EU H2020-funded Research Infrastructure Grant “European Virus Archive Global (EVAg)” (H2020—grant agreement no. 653316).

## REFERENCES

1. Aréchiga Ceballos N, Vázquez Morón S, Berciano JM, Nicolás O, Aznar López C, Juste J, Rodríguez Nevado C, Aguilar Setién A, Echevarría JE. 2013. Novel lyssavirus in bat, Spain. *Emerg Infect Dis* 19:793–795. <https://doi.org/10.3201/eid1905.121071>.
2. Gunawardena PS, Marston DA, Ellis RJ, Wise EL, Karawita AC, Breed AC, McElhinney LM, Johnson N, Banyard AC, Fooks AR. 2016. Lyssavirus in Indian flying foxes, Sri Lanka. *Emerg Infect Dis* 22:1456–1459. <https://doi.org/10.3201/eid2208.151986>.
3. Marston DA, McElhinney LM, Ellis RJ, Horton DL, Wise EL, Leech SL, David D, de Lamballerie X, Fooks AR. 2013. Next generation sequencing of viral RNA genomes. *BMC Genomics* 14:444. <https://doi.org/10.1186/1471-2164-14-444>.
4. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
5. Hunt M, Gall A, Ong SH, Brener J, Ferns B, Goulder P, Nastouli E, Keane JA, Kellam P, Otto TD. 2015. IVA: accurate *de novo* assembly of RNA virus genomes. *Bioinformatics* 31:2374–2376. <https://doi.org/10.1093/bioinformatics/btv120>.
6. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
7. Kuzmin IV, Wu X, Tordo N, Rupprecht CE. 2008. Complete genomes of Aravan, Khujand, Irkut and west Caucasian bat viruses, with special attention to the polymerase gene and non-coding regions. *Virus Res* 136:81–90. <https://doi.org/10.1016/j.virusres.2008.04.021>.