

# Genome Sequences of Nine Vesicular Stomatitis Virus Isolates from South America

Veronica L. Fowler,<sup>a</sup> David J. King,<sup>a</sup> Emma L. A. Howson,<sup>a</sup> Mikidache Madi,<sup>a</sup> Steven J. Pauszek,<sup>b</sup> Luis L. Rodriguez,<sup>b</sup> Nick J. Knowles,<sup>a</sup> Valérie Mioulet,<sup>a</sup> Donald P. King<sup>a</sup>

The Pirbright Institute, Pirbright, Woking, Surrey, United Kingdom<sup>a</sup>; Plum Island Animal Disease Center, Orient, New York, USA<sup>b</sup>

V.L.F. and D.J.K. contributed equally to this work.

**We report nine full-genome sequences of vesicular stomatitis virus obtained by Illumina next-generation sequencing of RNA, isolated from either cattle epithelial suspensions or cell culture supernatants. Seven of these viral genomes belonged to the New Jersey serotype/species (clade III), while two isolates belonged to the Indiana serotype/species.**

Received 22 February 2016 Accepted 29 February 2016 Published 14 April 2016

**Citation** Fowler VL, King DJ, Howson ELA, Madi M, Pauszek SJ, Rodriguez LL, Knowles NJ, Mioulet V, King DP. 2016. Genome sequences of nine vesicular stomatitis virus isolates from South America. *Genome Announc* 4(2):e00249-16. doi:10.1128/genomeA.00249-16.

**Copyright** © 2016 Fowler 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 Veronica L. Fowler, [veronica.fowler@pirbright.ac.uk](mailto:veronica.fowler@pirbright.ac.uk).

The genus *Vesiculovirus* (order *Mononegavirales*, family *Rhabdoviridae*) consists of nine virus species, four of which may cause vesicular stomatitis (VS), vesicular stomatitis New Jersey virus (VSNJV), vesicular stomatitis Indiana virus (VSIV), vesicular stomatitis Alagoas virus (VSAV), and Cocal virus (COCV). VS is endemic in southern Mexico, Central America, and the northern part of South America and mainly affects horses and cattle (1). Occasionally, pigs, sheep, goats, llamas, and alpacas may also be affected (1). VSV is a single-stranded negative-sense RNA virus of approximately 11 kb in length that encodes five proteins: nucleocapsid (N), phosphoprotein (P), matrix (M), glycoprotein (G), and polymerase (L) (2).

In this study, nine VS viruses were subjected to full-genome sequencing to be used for validation of new molecular diagnostic tests (Fowler et al., unpublished data). RNA was extracted using QIAamp spin column purification, and genomic DNA was depleted using DNase (Life technologies). First- and second-strand synthesis and DNA library generation was performed as previously described (3) but substituting a VSV-gene junction primer (5' AAA CTA ACA GAT ATC ATG GAC AG 3') for the first-strand synthesis stage. Consensus sequences were generated using *de novo* assembly; trimmed FASTQ files were processed using IDBA\_UD version 1.1.1, with an optimum *k*-mer length determined within the program (4). The identity of the contigs produced was confirmed via a BLAST search, and, if required, contigs were assembled manually using BioEdit version 7.2.5 (5). Using the *de novo* output as a reference, alignments were performed using BWA-MEM, SAM/BAM processing was performed using SAMtools (6), coverage plots were generated using BEDTools (7), and alignments were visually checked using Tablet (8). Final consensus sequences were generated using SAMtools mpileup. The total genome recovered for the seven New Jersey isolates ranged from 10,957 to 11,118 nucleotides in length with a mean coverage of 95.24 reads/genome position, obtaining 99.50% of the genomes. For these isolates, the total number of missing nucleotides are located in noncoding regions ( $n = 177$ ), N ( $n = 19$ ); Ecu-

dor/85  $n = 1$ , COL/1/93  $n = 4$ , 29453/COL/00  $n = 14$ ), M ( $n = 1$ ; 27405/COL/98), and L ( $n = 64$ ; COL/1/93  $n = 36$ , 27775/COL/98  $n = 1$ , 27405/COL/98  $n = 27$ ). These isolates grouped into VSNJV clade III, which contains all previously reported VSNJVs from South America. The total genome recovered for the two VSIV isolates (from serotype Indiana) ranged from 11,099 to 11,134 nucleotides in length with a mean coverage of 339.46 reads/genome site, obtaining 99.67% of the genomes. For these isolates, the total number of missing nucleotides are located in the non-coding region ( $n = 99$ ) and L ( $n = 2$ ; 29705/COL/01).

**Nucleotide sequence accession numbers.** The complete genome sequences for these isolates have been deposited in the GenBank database under the accession numbers [KU296051](https://www.ncbi.nlm.nih.gov/nuccore/KU296051) to [KU296059](https://www.ncbi.nlm.nih.gov/nuccore/KU296059).

## ACKNOWLEDGMENTS

We thank Alfonso Clavijo (Institute for Infectious Animal Diseases), the NCFAD-Canadian Food Inspection Agency, and colleagues in the WRLFMD (The Pirbright Institute, United Kingdom) for providing of virus samples. We would also like to thank Graham Freimanis for guidance with the adaptation of the protocol of Logan et al. (4).

## FUNDING INFORMATION

This work, including the efforts of David J. King, Emma L. A. Howson, Mikidache Madi, Nick J. Knowles, and Valerie Mioulet, was funded by BBSRC. This work, including the efforts of Veronica L. Fowler, Steven J. Pauszek, Luis Rodriguez, and Donald King, was funded by the Department for Homeland Security (HSHQDC-12-J-00415).

This work was funded by the U.S. Department of Homeland Security under award no. HSHQDC-12-J-00415 through Texas A&M AgriLife Research and the Institute for Infectious Diseases (IIAD). The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Additional core funding for sequencing work was provided by an Institute Strategic Pro-

gramme Grant from the Biotechnology and Biological Sciences Research Council (BBSRC).

## REFERENCES

- Rodríguez LL, Nichol ST. 1999. Vesicular stomatitis viruses, p. 1910–1919. In Webster RG, Granoff A (ed). *Encyclopedia of virology*, 2nd ed. Academic Press, London.
- Rodríguez LL, Pauszek SJ, Bunch TA, Schumann KR. 2002. Full-length genome analysis of natural isolates of vesicular stomatitis virus (Indiana 1 serotype) from North, Central and South America. *J Gen Virol* 83: 2475–2483. <http://dx.doi.org/10.1099/0022-1317-83-10-2475>.
- 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. <http://dx.doi.org/10.1093/bioinformatics/bts174>.
- Logan G, Freimanis GL, King DJ, Valdazo-González B, Bachanek-Bankowska K, Sanderson ND, Knowles NJ, King DP, Cottam EM. 2014. A universal protocol to generate consensus level genome sequences for foot-and-mouth disease virus and other positive-sense polyadenylated RNA viruses using the Illumina MiSeq. *BMC Genomics* 15:828. <http://dx.doi.org/10.1186/1471-2164-15-828>.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment Editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98.
- 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 (SAM) format and SAMtools. *Bioinformatics* 25:2078–2079. <http://dx.doi.org/10.1093/bioinformatics/btp352>.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. <http://dx.doi.org/10.1093/bioinformatics/btq033>.
- Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D. 2013. Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform* 14:193–202. <http://dx.doi.org/10.1093/bib/bbs012>.