Complete Genome Sequence of Rift Valley Fever Virus Strain Lunyo

Sarah Lumley,[a,b,c] Daniel L. Horton,[b] Denise A. Marston,[c] Nicholas Johnson,[c] Richard J. Ellis,[d] Anthony R. Fooks,[c,e] Roger Hewson[a,f]

Virology and Pathogenesis Group, Microbiology Services Division, Public Health England, Wiltshire, United Kingdom[a]; Department of Pathology and Infectious Disease, School of Veterinary Medicine, University of Surrey, Guildford, United Kingdom[b]; Wildlife Zoonoses and Vector-Borne Diseases Research Group, Animal and Plant Health Agency, Addlestone, Surrey, United Kingdom[c]; Specialist Scientific Support Department, Animal and Plant Health Agency, Addlestone, Surrey, United Kingdom[d]; Department of Clinical Infection, Microbiology and Immunology, Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom[e]; National Institute for Health Research, Health Protection Research Unit in Emerging and Zoonotic Infections, University of Liverpool, Liverpool, United Kingdom[f]

Using next-generation sequencing technologies, the first complete genome sequence of Rift Valley fever virus strain Lunyo is reported here. Originally reported as an attenuated antigenic variant strain from Uganda, genomic sequence analysis shows that Lunyo clusters together with other Ugandan isolates.

Rift Valley fever virus (RVFV) is a *Phlebovirus*, in the family *Bunyaviridae*, and was first isolated in 1930. It primarily affects ruminants, causing abortions and mortality, and in humans it causes a febrile illness, which is severe in 1% of cases (1). The virus cycles between its mosquito vector and mammalian hosts. The strain Lunyo was isolated from a pool of *Aedes* mosquitoes in the Lunyo Forest, Uganda, in 1955 (2). It was described as a variant strain due to varying reports of hemagglutinin and antibody neutralization (2–5). Pathogenicity in mice was initially low but increased with serial passage (2, 3), except in some strains of mice, hamsters, gerbils, and rats, where the virus remains attenuated (5, 6). Genome characterization is essential to improve understanding of the role sequence variation has on virulence.

Lunyo mouse brain suspension ($P \geq 11$) was passaged once in Vero E6 cells. Viral RNA was extracted from supernatant using the QIAamp viral RNA minikit (Qiagen). Double-stranded (ds) cDNA was synthesized from 50 ng of RNA using the cDNA Synthesis System (Roche) in accordance with the manufacturer's instructions. ds-cDNA was purified using Agencourt AMPure XP system (Beckman Coulter). A sequencing library was prepared from 1 ng of DNA using the Nextera XT-DNA library preparation kit (Illumina) according to the manufacturer's instructions. Sequencing was performed following standard Illumina protocols on the Illumina MiSeq with 150-bp paired-end reads. Total reads (8,556,470) were mapped to reference sequences (EU312121, DQ380217, DQ375419) using BWA version 0.7.5a-r405 (7) and visualized in Tablet (8). Consensus sequence was generated in modified samtools/vcfutils (9), and intermediate sequence was used as a reference for subsequent iteration mapping (10). A total of 67,991 reads were mapped (0.79%), resulting in near-complete small (S) (18,889 reads), complete medium (M) (22,833), and large (L) (26,269) segments with an average read depth of 1,208, 623, and 443, respectively. The S segment intergenic region nucleotides were resolved by amplicon-based next-generation sequencing using primers binding to terminal regions (11).

Phylogenetic analyses of the S, M, and L segments were performed in MEGA6. All complete RVFV sequences in GenBank ($n = 160$, 99, and 94 for the S, M and L segments, respectively) were aligned by ClustalW, and a maximum likelihood phylogenetic tree was constructed for each genome segment. Analysis of novel sequence data showed that all three Lunyo segments cluster with strains originating from Uganda, within the previously defined lineage E (12). Genetic similarity to a closely related strain isolated from mosquitoes in Entebbe, Uganda, in 1944 (DQ380156, DQ380191, DQ375429) demonstrated 97.5%, 98.6%, and 98.3% nucleotide identities in the S, M, and L segments, respectively. Results show 4 nonsynonymous substitutions in the nucleoprotein, 3 in the nonstructural protein, 8 in the glycoproteins, and 8 in the L protein. A single-nucleotide insertion in the intergenic region of the S segment conserves the segment size range of 1,690 to 1,962 (12). The contribution of a further complete genome to RVFV investigations will improve understanding of viral evolution and host interactions.

Nucleotide sequence accession numbers. Complete genomic sequences of RVFV strain Lunyo S, M, and L segments have been deposited in GenBank under accession numbers KU167025 to KU167027.

REFERENCES

1. **Mansfield KL, Banyard AC, McElhinney L, Johnson N, Horton DL, Hernández-Triana LM, Fooks AR**. 2015. Rift Valley fever virus: a review of diagnosis and vaccination, and implications for emergence in Europe. Vaccine **33**:5520–5531. http://dx.doi.org/10.1016/j.vaccine.2015.08.020.

2. **Weinbren MP, Williams MC, Haddow AJ**. 1957. A variant of Rift Valley fever virus. S Afr Med J **31**:951–957.

3. **Tomori O, Kasali O**. 1979. Pathogenicity of different strains of Rift Valley fever virus in Swiss albino mice. Br J Exp Pathol **60**:417–422.

4. **Swanepoel R, Struthers JK, Erasmus MJ, Shepherd SP, McGillivray GM, Shepherd AJ, Hummitzsch DE, Erasmus BJ, Barnard BJ**. 1986. Comparative pathogenicity and antigenic cross-reactivity of Rift Valley fever and other African phleboviruses in sheep. J Hyg (Lond) **97**:331–346.

5. **Peters CJ, Anderson GW, Jr**. 1981. Pathogenesis of Rift Valley fever. Contrib Epidemiol Biostat **3**:21–41.

6. **Battles JK, Dalrymple JM**. 1988. Genetic variation among geographic isolates of Rift Valley fever virus. Am J Trop Med Hyg **39**:617–631.

7. **Li H, Durbin R**. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics **26**:589–595. http://dx.doi.org/10.1093/bioinformatics/btp698.

8. **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.

9. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup 1000 Genome Project Data Processing**. 2009. The sequence alignment/map format and SAMtools. Bioinformatics **25**:2078–2079.

10. **Marston DA, Vázquez-Morón S, Ellis RJ, Wise EL, McElhinney LM, de Lamballerie X, Fooks AR, Echevarría JE**. 2015. Complete genomic sequence of European Bat lyssavirus 1, isolated from *Eptesicus isabellinus* in Spain. Genome Announc **3**(1):e01518-14. http://dx.doi.org/10.1128/genomeA.01518-14.

11. **Weingartl HM, Zhang S, Marszal P, McGreevy A, Burton L, Wilson WC**. 2014. Rift Valley fever virus incorporates the 78 kDa glycoprotein into virions matured in mosquito C6/36 cells. PLoS One **9**:e87385. http://dx.doi.org/10.1371/journal.pone.0087385.

12. **Bird BH, Khristova ML, Rollin PE, Ksiazek TG, Nichol ST**. 2007. Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry. J Virol **81**:2805–2816. http://dx.doi.org/10.1128/JVI.02095-06.