Complete Genome Sequences of Spondweni Viruses Isolated between 1958 and 1960

Petrus Jansen van Vuren,a Joe Kgaladi,a Venessa Patharoo,a Janusz T. Paweskaa

aCentre for Emerging Zoonotic and Parasitic Diseases, National Institute for Communicable Diseases, National Health Laboratory Service, Sandringham, South Africa

ABSTRACT Here, we report the complete genome sequences of 14 Spondweni viruses isolated in South Africa and Mozambique between 1958 and 1960. The sequences comprise 13 mosquito isolates and 1 human isolate following a documented laboratory infection. This study expands the publicly available data for this neglected virus from 4 to 18 sequences.

The Spondweni serogroup in the Flavivirus genus and Flaviviridae family comprises the human pathogens Zika virus and Spondweni virus (SPOV) (1–5). Zika virus (ZIKV) was first isolated in Uganda in 1947 but was acknowledged as a pathogen with severe impacts on human health only when recent outbreaks highlighted its propensity to cause severe neurological complications (6, 7). ZIKV is sexually transmissible, and SPOV was detected in testes and semen at a low level in a mouse model (8–10).

Six human infections with SPOV have been documented since its discovery in 1952 in Nigeria (1–5). ZIKV and SPOV infections result in similar disease presentations, ranging from asymptomatic to a mild or moderate febrile illness (11). The initial misidentification of SPOV as a strain of ZIKV and the extensive serological cross-reaction between these viruses have likely resulted in the misinterpretation of a number of early serosurveillance studies in Africa (11).

Until recently, SPOV was confined to sub-Saharan Africa. However, a SPOV isolate from Culex quinquefasciatus mosquitoes in Haiti in 2016 confirms that its distribution has expanded well beyond Africa (12). SPOV was isolated from four species of culicine mosquitoes in South Africa and Mozambique between 1958 and 1960 (4). We determined the full-genome sequence of these very early isolates.

SPOV isolates initially obtained from suckling mice (4) (Table 1) were cultured in VeroE6 cells, and viral RNA was extracted from the clarified supernatants (Qiagen viral RNA minikit). Virus cDNA was prepared as described before (13). Sequencing libraries were prepared using the Nextera DNA library preparation kit as recommended by the manufacturer (Illumina, USA) and sequenced on the MiSeq Illumina platform. Random hexamer and adapter sequences were removed from the reads using Cutadapt v1.21 (14). Quality filtering was performed using Prinseq-lite v0.20.4 (15). Sequencing reads were aligned to a reference sequence (Spondweni virus SM6-V1; GenBank accession number DQ859064) to determine consensus SPOV genomes. Reads were aligned to the reference sequence using Bowtie 2 (16), duplicates were removed using Picard, and a new consensus was generated using custom scripts (https://github.com/jtladner/Scripts/tree/master/reference-based_assembly). The mean read length for the 14 samples ranged between 264 and 272 nucleotides (nt), and total reads ranged between 1,250,200 and 1,832,112 nt. The number of reads mapped to the reference per sample ranged between 1,935 and 18,483 nt. Annotation was done manually using the CLC Genomics Workbench v10.0.1. Pairwise distance calculations between Spondweni isolate sequences were performed using MEGA v7.0 software (17).

As expected, the 14 SPOV isolates were closely related to previously sequenced...
SPOV isolates. The overall pairwise nucleotide distance between SPOV isolates was 0.61% (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a nucleotide distance ranging from 0.76% to 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%. The sequence similarities between AR1168 and AR1266 were identical, as were the sequences of AR1084 and AR1086 (range, 0% to 2.31%). The nucleotide sequences of isolates AR1084 and AR1086 were identical, as were the sequences of AR1077 and AR1081 (Table 1). The sequence of AR3061 was most closely related to previously sequenced isolates, with a pairwise nucleotide distance ranging from 0.31% to 2.25%. The sequence of the recent Haiti isolate was the most divergent but was still closely related to all other isolates, with a pairwise nucleotide distance between 0.76% and 2.31%.
Hills SL, Rabe IB, Lambert A, Abercrombie J, Martin SW, Gould CV, Oussayef N, Polen KN, Kuehnert MJ, Pillai SK, Petersen LR, Honein MA, Jamieson DJ, Brooks JT. 2016. Update: interim guidance for preconception counseling and prevention of sexual transmission of Zika virus for persons with possible Zika virus exposure—United States, September 2016. MMWR Morb Mortal Wkly Rep 65:1077–1081. https://doi.org/10.15585/mmwr.mm6539e1.

10. McDonald EM, Duggal NK, Brault AC. 2017. Pathogenesis and sexual transmission of Spondweni and Zika viruses. PLoS Negl Trop Dis 11: e0005990. https://doi.org/10.1371/journal.pntd.0005990.

11. Haddow AD, Woodall JP. 2016. Distinguishing between Zika and Spondweni viruses. Bull World Health Organ 94:711–71A. https://doi.org/10.2471/BLT.16.181503.

12. White SK, Lednicky JA, Okech BA, Morris J, Dunford JC. 2018. Spondweni virus in field-caught Culex quinquefasciatus mosquitoes, Haiti, 2016. Emerg Infect Dis 24:1765–1767. https://doi.org/10.3201/eid2409.171957.

13. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Ghedin E, Spiro DJ. 2008. Viral genome sequencing by random priming methods. BMC Genomics 9:5. https://doi.org/10.1186/1471-2164-9-5.

14. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBO J 17:10–12. https://doi.org/10.14806/ej.17.1.200.

15. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomics datasets. Bioinformatics 27:863–864. https://doi.org/10.1093/bioinformatics/btr026.

16. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.

17. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33: 1870–1874. https://doi.org/10.1093/molbev/msw054.