genomeA
Journals.ASM.org

Complete Genomic Sequence for an Avian Group G Rotavirus from South Africa

Karla M. Stucker,[a] Timothy B. Stockwell,[a] Martin M. Nyaga,[b] Rebecca A. Halpin,[a] Nadia Fedorova,[a] Asmik Akopov,[a] Harry Ngoveni,[b] Ina Peenze,[b] Mapaseka L. Seheri,[b] M. Jeffrey Mphahlele,[b] David E. Wentworth[a]*

J. Craig Venter Institute, Rockville, Maryland, USA[a]; South African Medical Research Council, Diarrhoeal Pathogens Research Unit (MRC-DPRU), Faculty of Health Sciences, Sefako Makgatho Health Sciences University, MEDUNSA, Pretoria, South Africa[b]

* Present address: David E. Wentworth, Virology Surveillance and Diagnosis Branch, Influenza Division, Centers for Disease Control and Prevention, Atlanta, Georgia, USA.

We report the first complete sequence for an avian group G rotavirus (RVG) genome from Africa, which is the third publically available RVG genome. These RVG genomes are highly diverse, especially in their VP4, VP7, NSP4, and NSP3 segments, indicating that RVG diversity is comparable to that of rotavirus A.

Received 30 January 2015 **Accepted** 4 February 2015 **Published** 12 March 2015

Citation Stucker KM, Stockwell TB, Nyaga MM, Halpin RA, Fedorova N, Akopov A, Ngoveni H, Peenze I, Seheri ML, Mphahlele MJ, Wentworth DE. 2015. Complete genomic sequence for an avian group G rotavirus from South Africa. Genome Announc 3(2):e00107-15. doi:10.1128/genomeA.00107-15.

Address correspondence to David E. Wentworth, dwentworth@cdc.gov.

Rotaviruses cause gastrointestinal diseases in many vertebrate species worldwide. As members of the *Reoviridae* family, rotaviruses have double-stranded RNA (dsRNA) genomes that comprise 11 segments. Eight rotavirus groups/species—rotaviruses A through H—have been defined by their VP6 proteins (1). Here, we report the first complete avian group G rotavirus (RVG) genome from Africa.

The RVG genome was extracted and amplified from a semi-diarrheal fecal sample collected from an otherwise asymptomatic 5-month-old female chicken (*Gallus gallus* domesticus) in Nkomo Village, Giyani, South Africa in 2011 using a sequence-independent amplification technique (2). Briefly, nucleic acids were extracted using TRI Reagent LS, and the dsRNA genome was enriched using lithium chloride precipitation. The PC3-T7 loop adapter was ligated to the 3′ ends of the dsRNA segments using T4 RNA ligase, and the cDNA was PCR amplified using primer PC2 (2). The amplified genome segments underwent standard bar coding and library construction for sequencing on the Ion Torrent, as well as a separate sequence-independent single-primer amplification (SISPA) using bar-coded random hexamers (3, 4) for sequencing on Illumina MiSeq v2 and HiSeq 2000.

Sequencing reads from all platforms were sorted by bar code, and bar codes and random hexamers were trimmed prior to assembly. Reads from all sequencing platforms were combined and *de novo* assembled using CLC Bio's clc_novo_assemble program. Assemblies were performed before any RVG genomes were publically available; therefore, distant homology searches using tblastx and HMM methods were used to assign contigs to the correct rotavirus segments. Once initial *de novo* contigs were identified for each segment, iterative cycles of mapping assembly for each segment allowed them to be built out through their termini using CLC Bio's clc_ref_assemble_long program to identify additional reads and extend the termini of each segment, followed by *de novo* assembly using all mapped reads. The iterations halted when each of the segment contigs reached the canonical GG at the 5′ terminus and CC at the 3′ terminus (5), with the PC2 primer sequence immediately beyond the termini. For this genome, all segments shared a 9-bp conserved 3′ terminal sequence (5′-TAAAGACCC-3′). For two regions each in VP1 and VP2, primers were designed and Sanger sequencing performed to improve coverage and connect contigs. The final assemblies were annotated using VIGOR (6, 7) and submitted to GenBank.

Comparing this African RVG genome with the recently deposited GenBank reference RVG genome from Germany (accessions NC_021580 through NC_021590) (8) shows a high degree of sequence diversity, with nucleotide percent identities ranging from 55.58% for VP4 to 94.54% for NSP5 and amino acid percent identities ranging from 40.73% for VP4 to 98.21% for VP6. As is typical for rotavirus species, the VP4, VP7, NSP4, and NSP3 segments are the most diverse among the 3 available RVG genomes from South Africa, Germany, and Hong Kong (8, 9). These comparisons demonstrate that a large amount of RVG diversity exists in nature and much of it remains to be sequenced.

Nucleotide sequence accession numbers. The genomic sequence data are deposited in GenBank under the accession numbers KJ752079 through KJ752089.

ACKNOWLEDGMENTS

We thank Daniel H. Haft for his help building the hidden Markov models that were used to annotate some of the RVG segments, as well as the members of the Joint Technology Center at the J. Craig Venter Institute (JCVI) for performing the Illumina sequencing and Susmita Shrivastava in the JCVI Bioinformatics Group for submitting the genome to NCBI. We also thank members of the MRC/Diarrhoeal Pathogens Research Unit who assisted in generating the cDNA using a sequence-independent amplification technique, as well as P. J. Loock and his team from the Limpopo Veterinary Laboratory Services for collecting and providing the sample.

The data for the manuscript and its preparation were generated while D.E.W. was employed at JCVI. The opinions expressed in this article are the authors' own and do not reflect the views of the Centers for Disease

Control and Prevention, the Department of Health and Human Services, or the U.S. government.

This project was funded with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under contract number HHSN272200900007C (Genome Sequencing Center for Infectious Diseases). The project was also funded by the South Africa Medical Research Council and the Poliomyelitis Research Foundation (PRF) of South Africa.

REFERENCES

1. **Matthijnssens J, Otto PH, Ciarlet M, Desselberger U, Van Ranst M, Johne R.** 2012. VP6-sequence-based cutoff values as a criterion for rotavirus species demarcation. Arch Virol **157:**1177–1182. http://dx.doi.org/10.1007/s00705-012-1273-3.
2. **Potgieter AC, Page NA, Liebenberg J, Wright IM, Landt O, van Dijk AA.** 2009. Improved strategies for sequence-independent amplification and sequencing of viral double-stranded RNA genomes. J Gen Virol **90:**1423–1432. http://dx.doi.org/10.1099/vir.0.009381-0.
3. **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. http://dx.doi.org/10.1186/1471-2164-9-5.
4. **Djikeng A, Spiro D.** 2009. Advancing full length genome sequencing for human RNA viral pathogens. Future Virol **4:**47–53. http://dx.doi.org/10.2217/17460794.4.1.47.
5. **Trojnar E, Otto P, Roth B, Reetz J, Johne R.** 2010. The genome segments of a group D rotavirus possess group A-like conserved termini but encode group-specific proteins. J Virol **84:**10254–10265. http://dx.doi.org/10.1128/JVI.00332-10.
6. **Wang S, Sundaram JP, Spiro D.** 2010. VIGOR, an annotation program for small viral genomes. BMC Bioinformatics **11:**451. http://dx.doi.org/10.1186/1471-2105-11-451.
7. **Wang S, Sundaram JP, Stockwell TB.** 2012. VIGOR extended to annotate genomes for additional 12 different viruses. Nucleic Acids Res **40:**W186–W192. http://dx.doi.org/10.1093/nar/gks528.
8. **Kindler E, Trojnar E, Heckel G, Otto PH, Johne R.** 2013. Analysis of rotavirus species diversity and evolution including the newly determined full-length genome sequences of rotavirus F and G. Infect Genet Evol **14:**58–67. http://dx.doi.org/10.1016/j.meegid.2012.11.015.
9. **Phan TG, Vo NP, Boros Á, Pankovics P, Reuter G, Li OT, Wang C, Deng X, Poon LL, Delwart E.** 2013. The viruses of wild pigeon droppings. PLoS One **8:**e72787. http://dx.doi.org/10.1371/journal.pone.0072787.