





Genome Sequences of Novel Members of Previously Described DNA and RNA Virus Families, Isolated from Feces of a Drill Monkey in Nigeria

U. George,^a C. Simsek,^{b,c}  T. O. C. Faleye,^{d,e} O. Arowolo,^f A. Oragwa,^g  O. M. Adewumi,^{d,h} J. Matthijnsens,^{b,c} J. A. Adeniji^{d,h,i}

^aDepartment of Biological Sciences, Redeemer's University, Ede, Nigeria

^bDepartment of Microbiology, Immunology, and Transplantation, KU Leuven, Leuven, Belgium

^cLaboratory of Viral Metagenomics, Rega Institute, KU Leuven, Leuven, Belgium

^dDepartment of Virology, Faculty of Basic Medical Sciences, College of Medicine, University of Ibadan, Ibadan, Nigeria

^eCenter for Human Virology and Genomics, Microbiology Department, Nigerian Institute of Medical Research, Lagos, Nigeria

^fViral Vaccines Unit, National Veterinary Research Institute, Vom, Nigeria

^gDepartment of Veterinary Microbiology and Pathology, University of Jos, Jos, Nigeria

^hInfectious Disease Institute, College of Medicine, University of Ibadan, Ibadan, Nigeria

ⁱWHO National Polio Laboratory, University of Ibadan, Ibadan, Nigeria

ABSTRACT The genomes of seven novel members of previously described DNA and RNA virus families are described here. These viruses were recovered using a viral metagenomic approach from the stool of a drill monkey (*Mandrillus leucophaeus*) housed in a sanctuary in Cross River State, Nigeria.

Zoonotic transmission of viruses between nonhuman primates (NHPs) and humans and its impact on the emergence and reemergence of viral pathogens have been well documented (1–3). Despite such studies, we are not aware of any metagenomic exploration of virus diversity among NHPs in Nigeria. Hence, in an effort to explore virus diversity among NHPs in Nigeria, we performed a metagenomic survey of virus diversity in a stool sample from a drill monkey (*Mandrillus leucophaeus*) in a sanctuary in the South South region of the country. Drills are among Africa's endangered mammals (because their current population in the wild may be <10,000) and are listed on the high conservation priority list by the International Union for Conservation of Nature (4). Here, we describe the genomes of novel members of previously described DNA and RNA virus families in a stool sample from a drill monkey in Nigeria.

The fresh fecal sample analyzed in this study was collected from the floor of an encampment housing drill monkeys in the Afi Mountain Wildlife Sanctuary in Cross River State, Nigeria. The sample was resuspended (1:9) in phosphate-buffered saline and subsequently subjected to the NetoVIR protocol (5). Briefly, after homogenization and centrifugation at 17,000 × *g* for 3 min, the stool suspension was filtered (0.8- μ m filter), and the filtrate was treated with nucleases. The treated stool suspension was then subjected to nucleic acid extraction using a Qiagen total nucleic acid extraction kit, and both RNA and DNA were randomly amplified using the whole-transcriptome amplification kit 2 (Sigma-Aldrich). Library preparation was subsequently done using the Nextera XT DNA kit. The library was then subjected to paired-end sequencing (2 × 150 bp) using the NextSeq platform (Illumina). Trimming was done using Trimmomatic (6), while assembly was done using SPAdes (7) and MEGAHIT (8). DIAMOND (9) was used to annotate the contigs obtained using the NCBI nonredundant protein database as the reference. The output was displayed using KronaTools to facilitate viral contig identification. All software was used with default settings.

Citation George U, Simsek C, Faleye TOC, Arowolo O, Oragwa A, Adewumi OM, Matthijnsens J, Adeniji JA. 2020. Genome sequences of novel members of previously described DNA and RNA virus families, isolated from feces of a drill monkey in Nigeria. *Microbiol Resour Announc* 9:e00092-20. <https://doi.org/10.1128/MRA.00092-20>.

Editor John J. Dennehy, Queens College

Copyright © 2020 George 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 O. M. Adewumi, adewumi1@hotmail.com.

Received 29 January 2020

Accepted 3 April 2020

Published 23 April 2020

TABLE 1 Assembly and characterization details of the novel genomes described in this study

Family	Genus (genome type)	Species ^a	GenBank accession no.	Genome size (nucleotides)	Coverage (fold)	No. (%) of mapped reads ^b	GC content (%)	Novelty type	Criterion for novelty ^c	GenBank accession no. for most similar sequence(s) (% similarity) ^d	Likely host
<i>Picornaviridae</i>	<i>Enterovirus</i> (RNA)	<i>Enterovirus J</i>	MN265403	7,282	39.48	2,294 (0.03)	44	Serotype	<88% aa sequence identity in VP1	AHY21607.1 (64.04)	Mammal
<i>Picornaviridae</i>	<i>Sapelovirus</i> (RNA)	<i>Sapelovirus C^e</i>	MN784122	8,207	182.64	12,036 (0.14)	39	Species	<70% aa sequence identity in complete polyprotein	ACG63545 (59.15)	Mammal
<i>Picornaviridae</i>	<i>Cosavirus</i> (RNA)	<i>Cosavirus G^e</i>	MN784123	7,714	798.03	49,316 (0.59)	47	Species	<65% aa sequence identity in P1	FJ438907 (43.59)	Mammal
<i>Dicistroviridae</i>	<i>Aparavirus</i> (RNA)	ABPV complex	MN296283	9,533	187.16	14,189 (0.17)	39	Strain	Not yet established	NP_851403.1 (78.79 [NS]); NP_851404.2 (83.27 [SP])	Insect
<i>Parvoviridae</i>	<i>Protoparvovirus</i> (DNA)	<i>BokiPPVirus^e</i>	MN265404	4,310	94.59	3,285 (0.04)	42	Species	<85% aa sequence identity in NS1	QDI06029.1 (71.50)	Mammal
<i>Parvoviridae</i>	<i>Ambidensovirus</i> (DNA)	<i>BokiADVirus^e</i>	MN784124	5,260	58.99	2,529 (0.03)	37	Species	<85% aa sequence identity in NS1	NP_046813.1 (71.93)	Insect
<i>Circoviridae</i>	<i>Cyclovirus</i> (DNA)	<i>BokiCVirus^e</i>	MN784125	1,800	4,604.92	67,307 (0.8)	49	Species	<80% genome-wide nucleotide sequence identity	GQ404854 (67.82)	Mammal

^aABPV, acute bee paralysis virus; BokiPPVirus, Boki Af Mountain protoparvovirus; BokiADVirus, Boki Af Mountain ambidensovirus; BokiCVirus, Boki Af Mountain cyclovirus.

^bThe total number of reads generated was 8,416,538.

^caa, amino acid; NS, nonstructural protein; SP, structural protein. Assemblies were primarily done using SPAdes. MEGAHIT assembly was done to resolve unclear assemblies (genome organizations) recovered from SPAdes (e.g., SPAdes assembled *Sapelovirus C* with two open reading frames, while MEGAHIT assembled it with one; the two were of similar lengths, however).

^dAll similarity values are amino acid similarities (determined using BLASTp) except for the value for the *Cyclovirus* entry, which is nucleotide similarity (determined using BLASTn).

^eLikely new species.

Seven near-complete genomes of novel members of previously described DNA and RNA virus families were recovered (Table 1), representing 150,956 (1.79%) of the 8,416,538 reads generated. The three DNA viruses are members of the families *Parvoviridae* (*Protoparvovirus* and *Ambidensovirus*) and *Circoviridae* (*Cyclovirus*). Three of the four RNA viruses are members of the *Picornaviridae* (*Cosavirus*, *Enterovirus*, and *Sapellovirus*), while the fourth is a member of the *Dicistroviridae* (*Aparavirus*). Unlike the other five genomes, the *Aparavirus* (*Dicistroviridae*) and *Ambidensovirus* (*Parvoviridae*) representatives might be insect viruses and are most likely relics from the diet of the drills (Table 1).

The genomes of seven novel members of previously described DNA and RNA virus families found in the stool of a drill monkey in Cross River State, Nigeria, are described here. A number of novel partial genomes were also found (data not shown). Here, we show that, by conserving drills, we might unknowingly have also conserved a previously undescribed virome. Whether these viruses are pathogenic in drills or might have zoonotic potential remains to be investigated. Current efforts are directed at better exploring virus diversity in this conserved ecosystem using a noninvasive approach, and our preliminary data suggest that these viruses are circulating.

Data availability. The genomes described here have been deposited under SRA number [PRJNA599206](https://www.ncbi.nlm.nih.gov/sra/PRJNA599206) and GenBank accession numbers [MN265403](https://www.ncbi.nlm.nih.gov/nucl/MN265403), [MN265404](https://www.ncbi.nlm.nih.gov/nucl/MN265404), [MN296283](https://www.ncbi.nlm.nih.gov/nucl/MN296283), and [MN784122](https://www.ncbi.nlm.nih.gov/nucl/MN784122) to [MN784125](https://www.ncbi.nlm.nih.gov/nucl/MN784125).

ACKNOWLEDGMENTS

We thank the management and staff of the Drill Rehabilitation and Breeding Centre and the Pandrillus Foundation for allowing and helping us to collect the sample analyzed in this study. We are particularly grateful to Innocent Itakwu, Asuquo Ani, Nsikan Eniniket, Basse Agbo, and Irene Edem for help with sample collection. We thank the next-generation sequencing (NGS) team at KU Leuven for contributions in the preparation of NGS libraries and the generation of NGS data.

This study was not funded by any specific grant, but publication costs were defrayed via a TETFund grant to J.A.A.

REFERENCES

- World Health Organization. 2012. Research priorities for zoonoses and marginalized infections. World Health Organization, Geneva, Switzerland.
- Formenty P, Boesch C, Wyers M, Steiner C, Donati F, Dind F, Walker F, Le Guenno B. 1999. Ebola virus outbreak among wild chimpanzees living in a rain forest of Cote d'Ivoire. *J Infect Dis* 179(Suppl 1):S120–S126. <https://doi.org/10.1086/514296>.
- Locatelli S, Peeters M. 2012. Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. *AIDS* 26:659–673. <https://doi.org/10.1097/QAD.0b013e328350fb68>.
- Oates JF, Butynski TM. 2008. *Mandrillus leucophaeus*. IUCN Red List Threatened Species 2008:e.T12753A3377185. <https://doi.org/10.2305/IUCN.UK.2008.RLTS.T12753A3377185.en>.
- Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Van Ranst M, Heylen E, Matthijssens J. 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep* 5:16532. <https://doi.org/10.1038/srep16532>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
- Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.