



## Circular Single-Stranded DNA Virus (*Microviridae: Gokushovirinae: Jodiemicrovirus*) Associated with the Pathobiome of the Flat-Back Mud Crab, *Eurypanopeus depressus*

✉ Jamie Bojko,<sup>a</sup> Krista A. McCoy,<sup>b</sup> Donald C. Behringer,<sup>a</sup> April M. H. Blakeslee<sup>b</sup>

<sup>a</sup>Emerging Pathogens Institute, University of Florida, Gainesville, Florida, USA

<sup>b</sup>Biology Department, East Carolina University, Greenville, North Carolina, USA

**ABSTRACT** A single-stranded DNA (ssDNA) virus is presented from a metagenomic data set derived from *Alphaproteobacteria*-infected hepatopancreatic tissues of the crab *Eurypanopeus depressus*. The circular virus genome (4,768 bp) encodes 14 hypothetical proteins, some similar to other bacteriophages (*Microviridae*). Based on its relatedness to other *Microviridae*, this virus represents a member of a novel genus.

*Microviridae* is a viral family with two subfamilies and 6 genera (1). It contains viruses that infect prokaryotes exclusively. Metagenomic techniques have unearthed the diversity of this family through evaluations of environmental, culture stock, and animal specimens (1–3). Microviruses have not been identified from *Rickettsiales* endosymbionts of Crustacea. Other bacteriophages isolated from crustacean microbiomes/pathobiomes show potential for phage therapy, avoiding the overuse of antibiotics in aquaculture (4). Few models exist to easily test this scenario in the laboratory.

We obtained DNA using a Zymo kit (D4070) on homogenized hepatopancreatic tissues of *Eurypanopeus depressus* ( $n = 1$ ), a panopeid crab from meso- and euryhaline locations across the Gulf of Mexico and Atlantic North America. The specimen was collected from a euryhaline site in North Carolina (Hoop Pole Creek, Atlantic Beach) in December 2018. A total of 1  $\mu$ g of DNA was used to prepare a NEBNext Ultra DNA library for Illumina HiSeq (10 $\times$ ) sequencing (NEB, USA) with a PE150 cartridge. This resulted in 11 million reads (50 to 150 bp) that were assembled using SPAdes v.3.13.0 (using default parameters and k-mer lengths of 21, 33, 55, 77, 99, and 127) (5) from trimmed reads using Trimmomatic (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) (6). This resulted in 523,047 contigs (>500 bp) ( $N_{50}$ , 2,133;  $N_{75}$ , 1,340;  $L_{50}$ , 100,989;  $L_{75}$ , 211,827). The genome (4,768 bp) of a *Microviridae* sp. was identified based on high coverage (>1,000 $\times$ ), with a GC content of 33% and 14 hypothetical open reading frames (ORFs) (Fig. 1 and Table 1). The genome was annotated using ExpAsy (standard genetic code) (7) and GeneMarkS (virus) (8). The relatedness of the genes and their function was identified using BLASTP (E value < 10) and InterProScan (9). Phylogenetics were conducted using IQ-Tree (10) after MAFFT alignment (11) of the capsid protein (ORF-1). The virus is genetically related to the *Gokushovirinae* subfamily of the *Microviridae* and represents a basal member to the three genera *Bdellomicrovirus*, *Chlamydiamicrovirus*, and *Spiromicrovirus*, as well as multiple other undescribed isolates associated with bacterial endosymbionts of tortoises, marine invertebrates, and insects (Fig. 1).

Of the 14 hypothetical ORFs, 5 showed similarity to other proteins in GenBank (Table 1). The proteins included a major capsid protein, DNA pilot protein, and the replicator initiator protein, which showed 32 to 44% amino acid similarity to other

**Citation** Bojko J, McCoy KA, Behringer DC, Blakeslee AMH. 2019. Circular single-stranded DNA virus (*Microviridae: Gokushovirinae: Jodiemicrovirus*) associated with the pathobiome of the flat-back mud crab, *Eurypanopeus depressus*. Microbiol Resour Announc 8:e01026-19. <https://doi.org/10.1128/MRA.01026-19>.

**Editor** Jelle Matthijnsens, KU Leuven

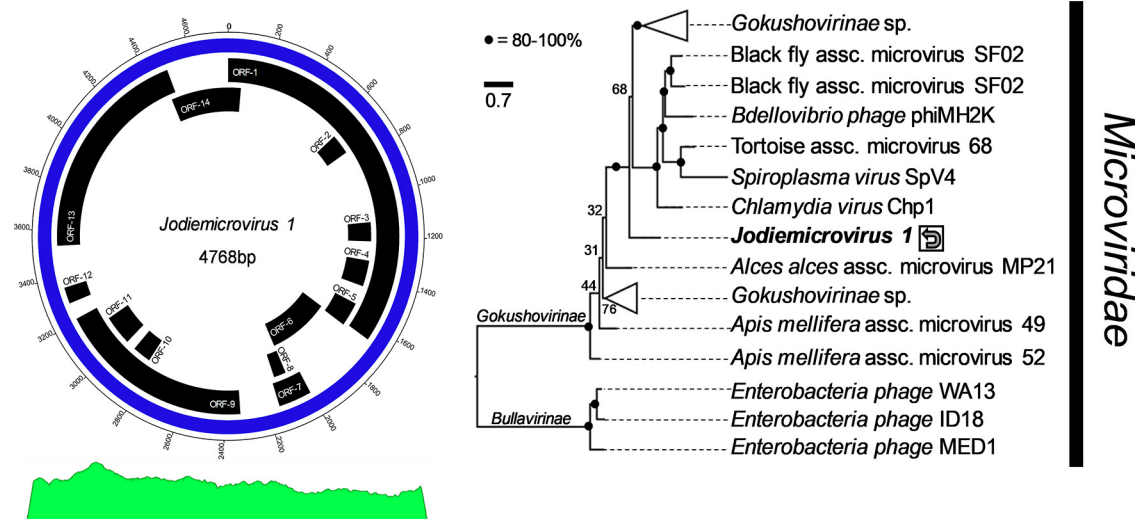
**Copyright** © 2019 Bojko et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Jamie Bojko, [jamie.bojko@ufl.edu](mailto:jamie.bojko@ufl.edu).

**Received** 4 September 2019

**Accepted** 22 October 2019

**Published** 21 November 2019



**FIG 1** Circular genome of *Jodiemicrovirus 1*, consisting of 4,768 bp, and phylogenetic comparison to other *Microviridae* using the capsid protein (ORF-1). The genome contains 14 hypothetical open reading frames. The chart at the bottom identifies the read coverage across the circular genome, representing 934,456 reads mapped to the genome, providing >1,000× coverage using CLC Genomics Workbench. The phylogenetic comparison included the MAFFT-aligned (11) capsid protein (716 positions) from multiple *Microviridae*. The maximum likelihood tree was inferred from 36 *Microviridae* spp. and was developed with the LG+F+G4 evolutionary model and 1,000 bootstraps in IQ-Tree (10). The final consensus tree (shown) had a log likelihood of -28,362.192 and scale of 0.7 units. The accession numbers used were [AXL15123](#), [AXQ65957](#), [QCS36953](#), [AXH77578](#), [AXL15643](#), [AZL82997](#), [AZL82921](#), [AXL14929](#), [YP\\_009218802](#), [AYQ58216](#), [AXL14945](#), [AZL82910](#), [AZL82729](#), [AZL83022](#), [AZL83017](#), [QCS36934](#), [QCS36961](#), [QCS37361](#), [AZL82956](#), [AZL82992](#), [AZL82871](#), [YP\\_009551424](#), [AZL82946](#), [QCQ84972](#), [AZL82926](#), [QCS37201](#), [QCQ84913](#), [AZL82717](#), [AZL82837](#), [YP\\_512416](#), [YP\\_512796](#), [AII27899](#), [NP\\_073538](#), [NP\\_044312](#), and [NP\\_598320](#).

*Microviridae* (Table 1). One virus was from an environmental sphagnum peat soil sample (12), and two viruses derived from the intestinal tract of *Ciona robusta* (Tunicata) (marine) (2). Two genes showed closest similarity to hypothetical bacterial genes (Table 1). Seven ORFs were identified internally to other ORFs, indicating the presence of putative overlapping genes recently discovered for the *Microviridae* (13). Based on its relatedness to known *Microviridae*, this genome might represent a novel genus (suggested, *Microviridae*: *Gokushovirinae*: *Jodiemicrovirus*).

To conclude, we present the genome of a bacteriophage likely to infect an undescribed member of the *Anaplasmataceae* which parasitizes the host hepatopancreas, identified via histology, electron microscopy, and genomics (our unpublished data). It may constitute a useful model system for understanding the effect of phage therapy relative to an intracellular bacterium causing disease in crustaceans.

**TABLE 1** Similarity and predicted function of the 14 hypothetical ORFs found in *Jodiemicrovirus 1*<sup>a</sup>

ORF	Predicted function	Closest hit (accession no.)	Similarity (%)	Coverage (%)	E value
1	Viral capsid	<i>Microviridae</i> sp. ( <a href="#">AXL15123</a> )	43.06	96	4e-138
2	Transmembrane	—	—	—	—
3	Signal peptide	<i>Acidimicrobiaceae</i> ( <a href="#">MBB33698</a> )	55.56	84	2.9
4	Unknown	—	—	—	—
5	Transmembrane	—	—	—	—
6	Unknown	—	—	—	—
7	Unknown	—	—	—	—
8	Unknown	—	—	—	—
9	DNA pilot protein	<i>Microviridae</i> sp. ( <a href="#">YP_009160339</a> )	33.33	40	1e-7
10	Unknown	—	—	—	—
11	Transmembrane	—	—	—	—
12	Signal peptide	—	—	—	—
13	Replication initiator protein	<i>Microviridae</i> sp. ( <a href="#">AXL15534</a> )	32.00	72	3e-28
14	Unknown	Bacteria ( <a href="#">EKD64965</a> )	40.38	41	6.6

<sup>a</sup> The coding orientation is positive in all cases. Data were analyzed using InterProScan and BLASTP. —, lack of significant protein similarity to any other known sequence data.

**Data availability.** The complete genome, annotation, and associated forward and reverse reads for this novel virus can be found under accession number [MN335165](#), BioProject number [PRJNA574411](#), and BioSample number [SAMN12567204](#).

## ACKNOWLEDGMENTS

We thank C. Moore and C. Erickson for their help in dissecting mud crabs for analysis. A.M.H.B. and K.A.M. thank East Carolina University for providing funds to support this research.

## REFERENCES

- Adriaenssens EM, Wittmann J, Kuhn JH, Turner D, Sullivan MB, Dutilh BE, Jang HB, van Zyl LJ, Klumpp J, Lobočka M, Moreno Switt AI, Rumnieks J, Edwards RA, Uchiyama J, Alfenas-Zerbini P, Petty NK, Kropinski AM, Barylski J, Gillis A, Clokie MRC, Prangishvili D, Lavigne R, Aziz R, Duffy S, Krupovic M, Poranen MM, Knezevic P, Enault F, Tong Y, Oksanen HM, Brister JR. 2018. Taxonomy of prokaryotic viruses: 2017 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Arch Virol* 163: 1125–1129. <https://doi.org/10.1007/s00705-018-3723-z>.
- Creasy A, Rosario K, Leigh B, Dishaw L, Breitbart M. 2018. Unprecedented diversity of ssDNA phages from the family *Microviridae* detected within the gut of a protochordate model organism (*Ciona robusta*). *Viruses* 10:404. <https://doi.org/10.3390/v10080404>.
- Walters M, Bawuro M, Christopher A, Knight A, Kraberger S, Stainton D, Chapman H, Varsani A. 2017. Novel single-stranded DNA virus genomes recovered from chimpanzee feces sampled from the Mambilla Plateau in Nigeria. *Genome Announc* 5:e01715-16. <https://doi.org/10.1128/genomeA.01715-16>.
- Lomeli-Ortega CO, Martínez-Díaz SF. 2014. Phage therapy against *Vibrio parahaemolyticus* infection in the whiteleg shrimp (*Litopenaeus vannamei*) larvae. *Aquaculture* 434:208–211. <https://doi.org/10.1016/j.aquaculture.2014.08.018>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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>.
- 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>.
- Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server, p 571–607. *In* Walker JM (ed), *The proteomics protocols handbook*. Humana Press, Totowa, NJ.
- Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29:2607–2618. <https://doi.org/10.1093/nar/29.12.2607>.
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Res* 33:W116–W120. <https://doi.org/10.1093/nar/gki442>.
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. 2016. IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* 44:W232–W235. <https://doi.org/10.1093/nar/gkw256>.
- Miller MA, Pfeiffer W, Schwartz T. 2011. The CIPRES science gateway: a community resource for phylogenetic analyses, p 41. *Proceedings of the 2011 TeraGrid Conference: extreme digital discovery*. Association for Computing Machinery, 18 to 21 July 2011, Salt Lake City, UT.
- Quaiser A, Dufresne A, Ballaud F, Roux S, Zivanovic Y, Colombet J, Sime-Ngando T, Francez A-J. 2015. Diversity and comparative genomics of *Microviridae* in Sphagnum-dominated peatlands. *Front Microbiol* 6:375. <https://doi.org/10.3389/fmicb.2015.00375>.
- Pavesi A. 2006. Origin and evolution of overlapping genes in the family *Microviridae*. *J Gen Virol* 87:1013–1017. <https://doi.org/10.1099/vir.0.81375-0>.