



Genome Sequences of Microviruses Identified in Gila Monster Feces

Courtney L. Collins,^a Dale F. DeNardo,^b Mellecha Blake,^b Jessica Norton,^a Kara Schmidlin,^a Rafaela S. Fontenele,^{a,b} Melissa A. Wilson,^{b,c} Simona Kraberger,^a  Arvind Varsani^{a,b,c,d}

^aThe Biodesign Center for Fundamental and Applied Microbiomics, Arizona State University, Tempe, Arizona, USA

^bSchool of Life Sciences, Arizona State University, Tempe, Arizona, USA

^cCenter for Evolution and Medicine, Arizona State University, Tempe, Arizona, USA

^dStructural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Observatory, Cape Town, South Africa

ABSTRACT The complete genome sequences of 33 microviruses were determined from fecal samples collected from 14 Arizona-dwelling Gila monsters using high-throughput sequencing. These microviruses with genomes 4,383 to 6,782 nucleotides (nt) long were broadly distributed across the 14 samples.

Gila monsters (*Heloderma suspectum*) are lizards found in the Sonoran Desert of North America (1). Little is known about the viruses associated with Gila monsters; to date, only adenoviruses (2, 3) and a genomovirus (4) have been reported. To further explore the viral diversity associated with Gila monsters, we analyzed the fecal samples collected directly from 14 individuals in Arizona in 2016. For each sample, 5 g of fecal material was homogenized in 20 ml of SM buffer (100 mM NaCl, 8 mM MgSO₄, 0.01% gelatin, and 50 mM Tris-HCl) and centrifuged at 10,000 × *g* for 10 min. The supernatant was first filtered through a 0.45- μ m filter followed by a 0.2- μ m syringe filter. Next, 10% (wt/vol) polyethylene glycol (PEG) was added to the filtrate and incubated overnight, and the viral particles were pelleted at 10,000 × *g* for 10 min. The pellet was resuspended in 500- μ l SM buffer, and 200 μ l was used to extract viral DNA using the High Pure viral nucleic acid kit (Roche Diagnostics, USA). Circular DNA in the extract was amplified using rolling circle amplification (RCA) with the Illustra TempliPhi kit (GE Healthcare, USA). The RCA products were used to generate (2 × 150-bp) Illumina libraries (individually barcoded) using the Hyper prep kit (Kapa Biosystems, USA) and multiplex sequenced on a lane of an Illumina NextSeq 500 sequencer at the Arizona State University (ASU) genomics core facility. Raw reads were quality trimmed with Trimmomatic v 0.39 (5) and *de novo* assembled using metaSPAdes v 3.12.0 (6). Contigs of >1,000 nucleotides (nt) were analyzed using VirSorter (7) for bacteriophage-like sequences (including microviruses). We identified 33 unique microvirus genomes across the 14 samples. They were determined to be circular based on terminal redundancy. Given that the same genome was identified in multiple samples, we mapped the reads derived from each fecal sample to the 33 unique microvirus genomes using BMap (8) to determine the distribution of microviruses across the samples using a threshold of 95% genome coverage for the purpose of this study (Fig. 1). All bioinformatic tools were run with default parameters.

Microviridae is a family of single-stranded DNA bacteriophages (9) that are found in a wide range of environments, such as seawater and animal gut samples (10–18). Microviruses have small, *T*=1, icosahedral capsids (9) and have two classified subfamilies, namely, *Gokushovirinae* and *Bullavirinae*. The 33 microvirus genome sequences (4,383 to 6,782 nt) identified in this study have GC contents of 29.9% to 55.4% (Fig. 1). The open reading frames were identified using RASTtk (19) and annotated based on BLASTP (20) similarities to proteins encoded by microvirus sequences available in GenBank. They all encode at least a

Citation Collins CL, DeNardo DF, Blake M, Norton J, Schmidlin K, Fontenele RS, Wilson MA, Kraberger S, Varsani A. 2021. Genome sequences of microviruses identified in Gila monster feces. *Microbiol Resour Announc* 10:e00163-21. <https://doi.org/10.1128/MRA.00163-21>.

Editor Kenneth M. Stedman, Portland State University

Copyright © 2021 Collins 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 Arvind Varsani, Arvind.varsani@asu.edu.

Received 11 February 2021

Accepted 25 February 2021

Published 18 March 2021

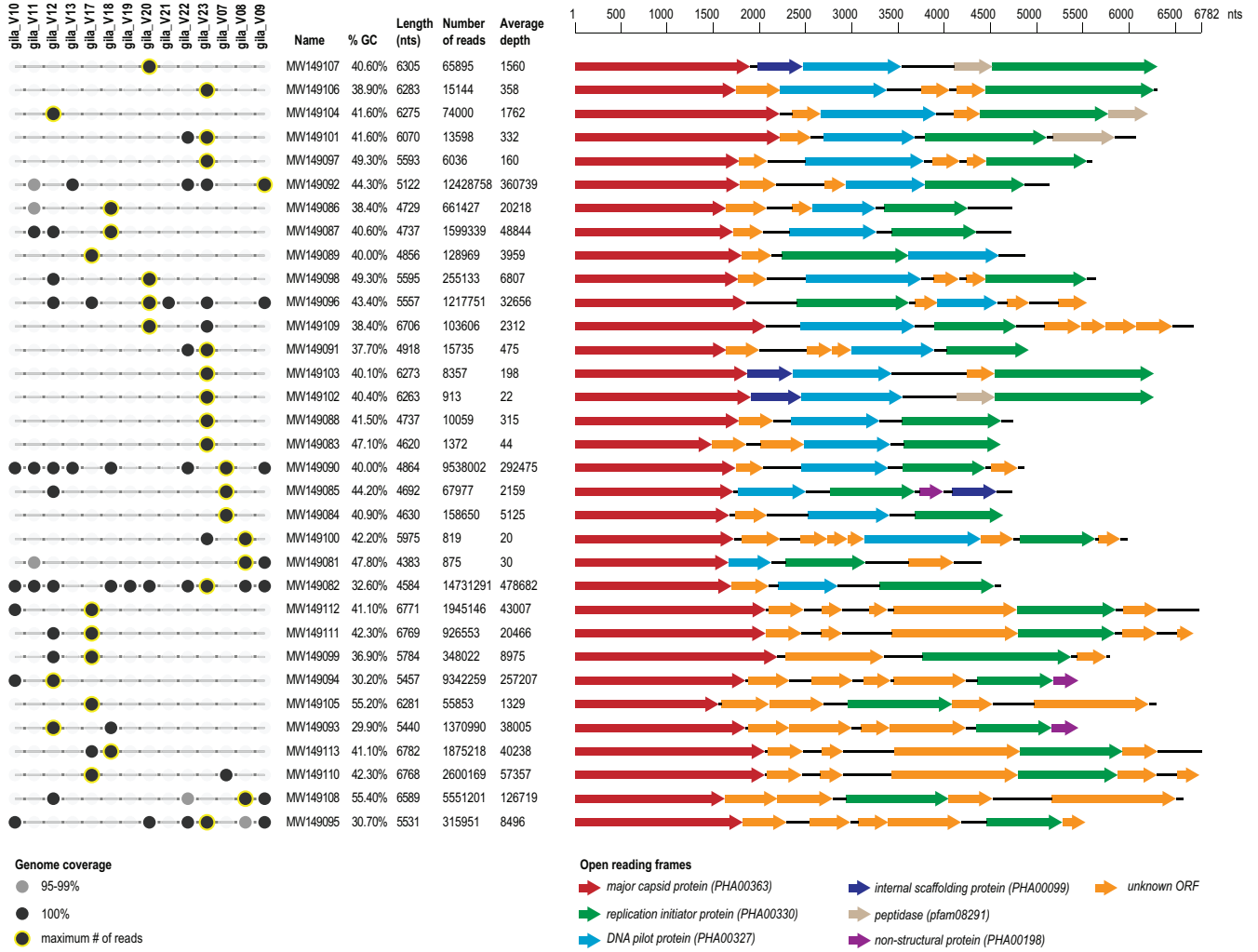


FIG 1 Summary of the 33 microvirus genomes identified in this study and their distribution across 14 Gila monster fecal samples. The solid dark-gray and black circles indicate 95% to 99% and 100% raw read genome coverage, respectively, per Gila monster fecal sample. The number of reads that mapped to the microvirus genome sequence and the depth of the read coverage are summarized, and the sample containing the highest number of reads for each microvirus is denoted by a black circle with a yellow outline. The genome organization is provided on the right with color-coded, open reading frames with details of putative protein families.

monocyte chemoattractant protein (MCP), a replication initiator protein, and their genomes have an average read fold depth ranging from 20 to 478,682 (Fig. 1). These microviruses have variable distribution across the samples (based on 95% genome coverage) ranging from 1 to 10 (Fig. 1). The MCP sequences, when analyzed with those of other microviruses available in GenBank (as of 7 December 2020), share pairwise identities in the range of 33.7% to 100% amino acid identity, as determined by SDT v 1.2 (21).

Data availability. The sequences of microviruses in this study have been deposited in the NCBI SRA database under project [PRJNA667500](#) and in GenBank under the accession numbers [MW149081](#) to [MW149113](#).

ACKNOWLEDGMENTS

This work was supported by funds from the Bidesign Institute and School of Life Sciences at Arizona State University awarded to A.V.

REFERENCES

1. Beck DD, Martin BE, Lowe CH. 2005. Biology of Gila monsters and bearded lizards. University of California Press, Berkeley, CA.
2. Papp T, Fledelius B, Schmidt V, Kajan G, Marschang R. 2009. PCR-sequence characterization of new adenoviruses found in reptiles and the first

- successful isolation of a lizard adenovirus. *Vet Microbiol* 134:233–240. <https://doi.org/10.1016/j.vetmic.2008.08.003>.
3. Wellehan JFX, Johnson AJ, Harrach B, Benkő M, Pessier AP, Johnson CM, Garner MM, Childress A, Jacobson ER. 2004. Detection and analysis of six lizard adenoviruses by consensus primer PCR provides further evidence of a reptilian origin for the adenoviruses. *J Virol* 78:13366–13369. <https://doi.org/10.1128/JVI.78.23.13366-13369.2004>.
 4. Somayaji V, DeNardo D, Wilson Sayres MA, Blake M, Waits K, Fontenele RS, Kraberger S, Varsani A. 2018. Genome sequence of a single-stranded DNA virus identified in Gila monster feces. *Microbiol Resour Announc* 7:e00925-18. <https://doi.org/10.1128/MRA.00925-18>.
 5. 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>.
 6. 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>.
 7. Roux S, Enault F, Hurwitz BL, Sullivan MB. 2015. VirSorter: mining viral signal from microbial genomic data. *PeerJ* 3:e985. <https://doi.org/10.7717/peerj.985>.
 8. Bushnell B. 2014. BMAP: a fast, accurate, splice-aware aligner. Lawrence Berkeley National Lab. Berkeley, CA.
 9. Cherwa JE, Fane BA. 2011. *Microviridae: microviruses and gokushoviruses*. eLS. John Wiley & Sons, Ltd, Chichester, UK.
 10. Roux S, Krupovic M, Poulet A, Debroas D, Enault F. 2012. Evolution and diversity of the Microviridae viral family through a collection of 81 new complete genomes assembled from virome reads. *PLoS One* 7:e40418. <https://doi.org/10.1371/journal.pone.0040418>.
 11. Kirchberger PC, Ochman H. 2020. Resurrection of a global, metagenomically defined gokushovirus. *Elife* 9:e51599. <https://doi.org/10.7554/eLife.51599>.
 12. Schmidlin S, Kraberger S, Fontenele RS, De Martini F, Chouvenec T, Gile G, Varsani A. 2019. Genome sequences of microviruses associated with *Coptotermes formosanus*. *Microbiol Resour Announc* 8:e00185-19. <https://doi.org/10.1128/MRA.00185-19>.
 13. Roux S, Enault F, Robin A, Ravet V, Personnic S, Theil S, Colombet J, Sime-Ngando T, Debroas D. 2012. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS One* 7:e33641. <https://doi.org/10.1371/journal.pone.0033641>.
 14. Orton JP, Morales M, Fontenele RS, Schmidlin K, Kraberger S, Leavitt DJ, Webster TH, Wilson MA, Kusumi K, Dolby GA, Varsani A. 2020. Virus discovery in desert tortoise fecal samples: novel circular single-stranded DNA viruses. *Viruses* 12:143. <https://doi.org/10.3390/v12020143>.
 15. Kraberger S, Cook CN, Schmidlin K, Fontenele RS, Bautista J, Smith B, Varsani A. 2019. Diverse single-stranded DNA viruses associated with honey bees (*Apis mellifera*). *Infect Genet Evol* 71:179–188. <https://doi.org/10.1016/j.meegid.2019.03.024>.
 16. Fontenele RS, Lacorte C, Lamas NS, Schmidlin K, Varsani A, Ribeiro SG. 2019. Single stranded DNA viruses associated with capybara faeces sampled in Brazil. *Viruses* 11:710. <https://doi.org/10.3390/v11080710>.
 17. 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>.
 18. Creasy A, Rosario K, Leigh BA, Dishaw LJ, Breitbart M. 2018. Unprecedented diversity of ssDNA Phages from the family detected within the gut of a protochordate model organism (*Ciona robusta*). *Viruses* 10:404. <https://doi.org/10.3390/v10080404>.
 19. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>.
 20. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
 21. Muhire BM, Varsani A, Martin DP. 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9:e108277. <https://doi.org/10.1371/journal.pone.0108277>.