




Complete Genome Sequence of the *Gordonia* Bacteriophage BiggityBass

Cyril J. Versoza,^{a,b} Abigail A. Howell,^{a,c} Tanya Aftab,^a Madison Blanco,^a Akarshi Brar,^a Elaine Chaffee,^a Nicholas Howell,^{a,d} Willow Leach,^d Jackelyn Lobatos,^a Michael Luca,^{a,e} Meghna Maddineni,^e Ruchira Mirji,^a Corinne Mitra,^a Maria Strasser,^a Saige Munig,^a Zeel Patel,^a Minerva So,^a Makena Sy,^a Sarah Weiss,^a Christopher D. Herren,^f Martha Smith Caldas,^f  Susanne P. Pfeifer^{a,b,g}

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

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

^cBiodesign Institute, Arizona State University, Tempe, Arizona, USA

^dSchool of Mathematical and Statistical Sciences, Arizona State University, Tempe, Arizona, USA

^eSchool of Molecular Sciences, Arizona State University, Tempe, Arizona, USA

^fDivision of Biology, Kansas State University, Manhattan, Kansas, USA

^gCenter for Mechanisms of Evolution, Arizona State University, Tempe, Arizona, USA

Cyril J. Versoza and Abigail A. Howell contributed equally to the project. Author order was determined on the basis of seniority.

ABSTRACT Here, we characterized the complete genome of the *Siphoviridae* BiggityBass, a lytic subcluster DR bacteriophage infecting *Gordonia terrae* CAG3. Its 63.2-kb genome contains 84 protein-coding genes, of which 40 could be assigned a putative function. BiggityBass is related most closely to AnClar and Yago84 with 90.61% and 90.52% nucleotide identity, respectively.

Bacteriophages are one of the most abundant organisms on our planet (1), and yet, their diversity remains poorly characterized. As part of the Howard Hughes Medical Institute Science Education Alliance–Phage Hunters Advancing Genomics and Evolutionary Science (HHMI SEA-PHAGES) course-based undergraduate research experience, we characterized the complete genome sequence of BiggityBass, an obligatory lytic *Gordonia terrae* CAG3 bacteriophage.

BiggityBass was obtained by direct plating of a filtered soil sample collected in Manhattan, Kansas, at the top of a hill on the Konza Prairie Nature Trail (39.108056 N, 96.596944 W), purifying a well-isolated plaque, and amplifying the plaque in *Gordonia terrae* CAG3 (which was propagated in peptone-yeast extract-calcium [PYCa] broth at 30°C for 48 h with aeration in an orbital incubator at 250 rpm), following the procedures outlined in the SEA-PHAGES Discovery Guide (<https://seaphagesphagediscoveryguide.helpdocsonline.com/home>).

Copper electron microscopy grids were prepared using uranium acetate staining, which demonstrated that BiggityBass exhibits a *Siphoviridae* morphology, characterized by a long, noncontractile tail and an icosahedral capsid that contains the double stranded DNA (dsDNA) (Fig. 1). Genomic DNA was extracted using the Promega Wizard DNA clean-up kit (washing columns twice with 80% propanol, without RNase or DNase pretreatment), and a library was prepared using the NEBNext Ultra II FS kit and sequenced on an Illumina MiSeq instrument to more than 900-fold coverage. Next, 411,782 high-quality, 150-bp single-end raw sequencing reads were *de novo* assembled using Newbler v.2.9 (2), yielding a complete genome of length 63,202 bp with circularly permuted ends. The genome exhibits a GC content of 69.4%, which is similar to that of its host *Gordonia terrae* (67.8%). The assembly was checked for completeness, accuracy, and genome termini using Consed v.29.0 (3). All software was executed using default settings.

Following the HHMI SEA-PHAGES Bioinformatics Guide (<https://seaphagesbioinformatics.helpdocsonline.com/home>), 84 genes were predicted using GLIMMER v.3.0.2 (4) and GeneMark

Editor John J. Dennehy, Queens College CUNY

Copyright © 2022 Versoza 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 Susanne P. Pfeifer, susanne.pfeifer@asu.edu.

The authors declare no conflict of interest.

Received 7 June 2022

Accepted 18 July 2022

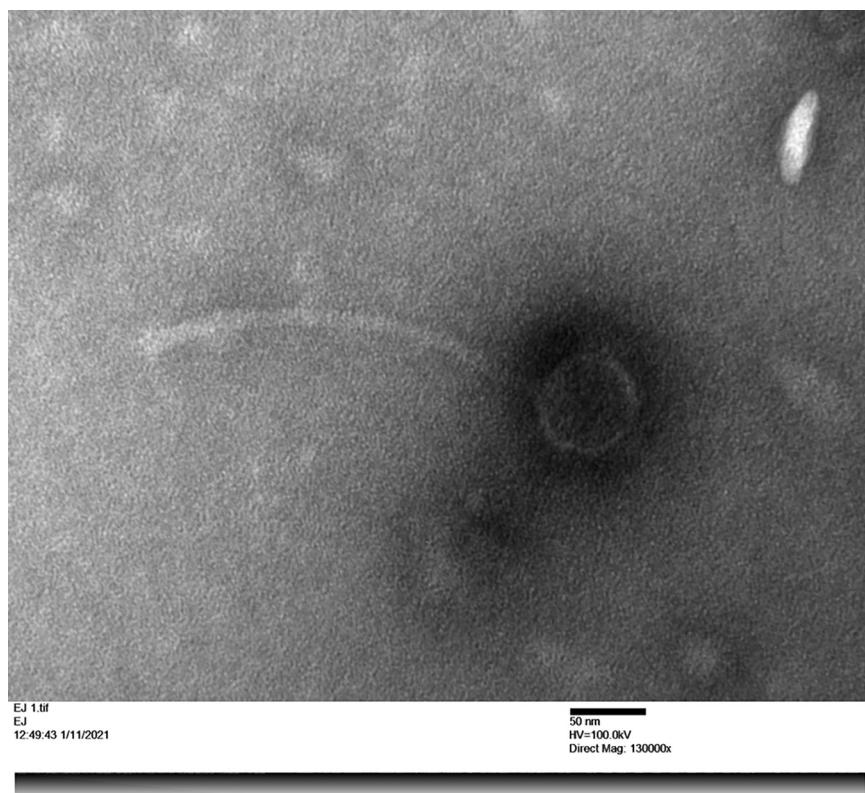


FIG 1 Transmission electron microscopy image showing BiggityBass *Siphoviridae* morphology, characterized by a long, noncontractile tail and an icosahedral capsid that contains the dsDNA.

v.2.5 (5) embedded within DNA Master (<http://cobamide2.bio.pitt.edu>), corresponding to a gene density of 1.31 genes/kb. Out of these 84 genes, 40 could be assigned a putative function using BLASTp v.2.13 (6) and HHPred (7) alignments against GenBank and the Protein Data Base, as well as Phamerator (8) to determine synteny with other cluster DR bacteriophages. Thereby, the position of the minor tail protein (gene 18) is of particular note, potentially hinting at a host range pattern that is different from other cluster DR bacteriophages. In addition, six genes were identified to be membrane proteins using TMHMM v.2.0 (9) and SOSUI v.1.11 (10). ARAGORN v.1.1 (embedded within DNA Master) and v.1.2.38 (11) as well as tRNAscan-SE v.2.0 (12) were used to search for tRNAs and transfer-messenger RNAs (tmRNAs), but none were found.

Multiple sequence alignments using MAFFT v.7 (13) indicated that BiggityBass is related most closely to AnClar (GenBank accession number [MN908693](#)) and Yago84 (GenBank accession number [MK801725](#)) (14) with a nucleotide sequence similarity of 90.61% and 90.52%, respectively.

Data availability. The whole-genome sequencing data are available through NCBI Sequence Read Archive (BioProject accession number [PRJNA488469](#); run number [SRR20167194](#)). The annotated genome assembly is available through NCBI GenBank under accession number [ON260813](#).

ACKNOWLEDGMENTS

This work was supported by a National Science Foundation CAREER grant to S.P.P. (DEB-2045343), Howard Hughes Medical Institute SEA-PHAGES program, and Arizona State University School of Life Sciences.

Bacteriophage isolation was performed at Kansas State University, Manhattan, KS; library preparation, sequencing, and *de novo* assembly were performed at the University of Pittsburgh, Pittsburgh, PA; and genome annotations and comparative analyses were performed at Arizona State University, Tempe, AZ.

We are grateful to Suhail Ghafoor for information technology (IT) support and Daniel A. Russell and Rebecca A. Garlena for *de novo* assembly, as well as Billy Biederman, Graham Hatfull, Deborah Jacobs-Sera, and Vic Sivanathan for training and continued support in the SEA-PHAGES program.

REFERENCES

1. Rohwer F. 2003. Global phage diversity. *Cell* 113:141. [https://doi.org/10.1016/s0092-8674\(03\)00276-9](https://doi.org/10.1016/s0092-8674(03)00276-9).
2. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380. <https://doi.org/10.1038/nature03959>.
3. Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res* 8:195–202. <https://doi.org/10.1101/gr.8.3.195>.
4. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 27:4636–4641. <https://doi.org/10.1093/nar/27.23.4636>.
5. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* 26:1107–1115. <https://doi.org/10.1093/nar/26.4.1107>.
6. 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).
7. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <https://doi.org/10.1093/nar/gki408>.
8. Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics* 12:395. <https://doi.org/10.1186/1471-2105-12-395>.
9. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580. <https://doi.org/10.1006/jmbi.2000.4315>.
10. Hirokawa T, Boon-Chieng S, Mitaku S. 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14:378–379. <https://doi.org/10.1093/bioinformatics/14.4.378>.
11. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. <https://doi.org/10.1093/nar/gkh152>.
12. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964. <https://doi.org/10.1093/nar/25.5.955>.
13. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
14. Pope WH, Butela KA, Garlena RA, Jacobs-Sera D, Russell DA, Warner MH, Hatfull GF, University of Pittsburgh SEA-PHAGES. 2020. Genome sequences of 20 bacteriophages isolated on *Gordonia terrae*. *Microbiol Resour Announc* 9: e01489-19. <https://doi.org/10.1128/MRA.01489-19>.