



Annotation of the Complete Genome Sequences of Bacteriophages Sara and Birdfeeder

Benjamin M. Adams,^a Joseph B. Adams,^a Reganne L. Brewster,^a Michael S. Cutler,^a Anthony E. Davis,^a Abby H. Gallegos,^a Jeremy S. Hernandez,^a Luke H. May,^a Emma G. Montoya,^a Andrew T. Reagan,^a Jack F. Shurley,^a Anna S. Grinath,^a  Michael A. Thomas^a

^aDepartment of Biological Sciences, Idaho State University, Pocatello, Idaho, USA

ABSTRACT Sara is a siphovirus with a linear 17,362bp genome containing 25 genes. Birdfeeder is a podovirus with a circularly permuted 53,897bp genome containing 52 genes. Sara and Birdfeeder were isolated from environmental samples in Plattsburgh, NY, USA and Forest Hill, MD, USA, respectively, using *Microbacterium foliorum* NRRL B-24224.

Characterizing bacteriophages improves understanding of the most plentiful biological constructs on earth and their human health applications (1, 2). As *Microbacterium* have been associated with plant drought resistance and meat spoilage, phages infecting this genus are of interest for the agricultural and food industries (3). Here, we describe two bacteriophages, Sara and Birdfeeder, that infect soil bacterium *M. foliorum*.

Phage isolation followed standard procedures (4). Sara was isolated from soil collected near the Saranac River in Plattsburgh, New York, (Global Positioning System [GPS] coordinates 44.69° N, 73.46° W), whereas Birdfeeder was isolated from soil collected underneath a bird feeder in Forest Hill, Maryland, (GPS coordinates 39.559167° N, 76.423611° W). Soil samples were washed with peptone/yeast/calcium (PYCa) liquid media and bacteriophages extracted through a 0.22- μ m filter. Filtrate was mixed with soft agar containing *M. foliorum* NRRL B-24224, overlaid on PYCa agar, and incubated at 20°C for 48 h. Sara produced small, clear pinpoint plaques while Birdfeeder produced large, haloed plaques. Both phages were purified with at least three rounds of plating.

DNA was extracted from high-titer lysates using the Promega Wizard DNA cleanup kit and prepared for sequencing using the NEBNext Ultrall FS kit. DNA was sequenced using Illumina MiSeq (v3 reagents), generating 375,319 and 131,281 150-bp unpaired reads for Sara and Birdfeeder, respectively. Raw reads were trimmed and assembled using Newbler v2.9 with default parameters, yielding a single contig for each phage genome; Consed v29 used used to check genomes for completeness and accuracy and to determine phage termini (5, 6). Genome characteristics are provided in Table 1. Based on gene content similarity of 35% of higher to phages in the Actinobacteriophage database, Sara and Birdfeeder are assigned to phage clusters EE and EK, respectively (7, 8). The GC content of the genomes are consistent with other members of their respective subclusters; additionally, the GC content for Sara is similar to that of the host bacteria, *M. foliorum* (68.7%) (3).

Phage genomes were auto-annotated using DNAmaster v5.23.6 (<http://cobamide2.bio.pitt.edu>) embedded with GeneMark v2.5 (9) and GLIMMER v3.02 (10), with start sites then refined using Phage Evidence Collection and Annotation Network (PECAAN) (<http://discover.kbrinsgd.org>), Phamerator v454 (11) and Starterator v1.0.1 (<https://seaphages.org/software>). A total of 52 and 25 protein-coding genes were identified in Birdfeeder and Sara, respectively (Table 1). Using Aragorn v1.2.41 (12) and tRNA-SE v2.0 (13), no tRNAs were detected in either phage genome. Putative gene functions were determined using BLAST v2.11.0 (14) and HHPred v2.0 (15). All software used default parameters.

Editor John J. Dennehy, Queens College CUNY

Copyright © 2022 Adams 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 Michael A. Thomas, mthomas@isu.edu.

The authors declare no conflict of interest.

Received 2 August 2022

Accepted 7 September 2022

Published 19 September 2022

TABLE 1 Genome characteristics for sara and birdfeeder genomes

Phage	Subcluster	Avg sequence coverage (x)	Genome size	Genome terminus arrangement	GC content	No. of protein coding genes
Sara	EE	3084	17,362bp	Linear, with 3' single-stranded overhangs (5'-CCCCCCCCA-3')	68.5%	25
Birdfeeder	EK	349	53,897bp	Circular	60.0%	52

For Sara, structure and assembly genes span the first two-thirds of the genome. With the exception of three genes (SEA_SARA_20 to SEA_SARA_22) encoding DNA-binding proteins, all genes for Sara are transcribed rightwards. In contrast, the first third of the Birdfeeder genome encodes for several DNA metabolism genes are transcribed leftwards, followed by all rightwards-transcribed genes that include structure and assembly genes and a gene that is notably 13,479 bp-long. No immunity repressor or integrase functions could be identified for either phage. For Sara, this is consistent with its clear plaque morphology and with phages in cluster EE consisting of lytic siphoviruses. The lifecycle of phages in cluster EK is unknown.

Data availability. GenBank Accession and Sequence Read Archive (SRA) numbers are [ON260812](https://doi.org/10.1093/bioinformatics/btt515) & [SRX14485116](https://doi.org/10.1093/bioinformatics/btt515), respectively (Sara) and [ON456346](https://doi.org/10.1093/bioinformatics/btt515) & [SRX14989441](https://doi.org/10.1093/bioinformatics/btt515), respectively (Birdfeeder).

ACKNOWLEDGMENTS

This study was kindly supported by the HHMI-sponsored SEA-PHAGES program, and the Idaho State University Department of Biological Sciences.

We thank Mallory Brownell, Temitayo Adegbite, and colleagues at State University of New York-Plattsburgh for the discovery and isolation of Sara, and C.J. Jednorski, E.N. Walters, J.R. Lebrasseur, and colleagues at Harford Community College for the discovery and isolation of Birdfeeder.

REFERENCES

- Aranaga C, Pantoja LD, Martínez EA, Falco A. 2022. Phage therapy in the era of multidrug resistance in bacteria: a systematic review. *Int J Mol Sci* 23:4577. <https://doi.org/10.3390/ijms23094577>.
- Principi N, Silvestri E, Esposito S. 2019. Advantages and limitations of bacteriophages for the treatment of bacterial infections. *Front Pharmacol* 10:513. <https://doi.org/10.3389/fphar.2019.00513>.
- Russell DA, Garlena RA, Hatfull GF. 2019. Complete genome sequence of *Microbacterium foliorum* NRRL B-24224, a host for bacteriophage discovery. *Microbiol Resour Announc* 8:e01467-18. <https://doi.org/10.1128/MRA.01467-18>.
- Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull GF. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD, USA. <https://seaphagesphagediscoveryguide.helpdocsonline.com>.
- Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes. *Methods Mol Biol* 1681:109–125. https://doi.org/10.1007/978-1-4939-7343-9_9.
- Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29:2936–2937. <https://doi.org/10.1093/bioinformatics/btt515>.
- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.
- Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, Russell DA, Warner MH, Hatfull GF, Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES). 2017. Bacteriophages of *Gordonia* spp. display a spectrum of diversity and genetic relationships. *mBio* 8:e01069-17. <https://doi.org/10.1128/mBio.01069-17>.
- Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* 33:W451–W454. <https://doi.org/10.1093/nar/gki487>.
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673–679. <https://doi.org/10.1093/bioinformatics/btm009>.
- 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>.
- 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>.
- 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>.
- 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).
- 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>.