




Complete Genome Sequence of *Microbacterium* Bacteriophage Erla

Mark Milhaven,^a Erin Hastings,^b Danielle Brister,^a Leo Cevallos,^a Sriya Chilukuri,^a Arshia Diyya,^a Aman Garg,^a Rosaura Hernandez,^a Daniel Kelly,^a Karina Lazo,^a Jennifer Le,^a Garrett Maag,^a Palak D. Marfatia,^a Rithik Mehta,^a Aram Nejad,^{a,c} Jade Porche,^a Alexander Queiroz,^a Daniel Sackett,^a Pablo Santos Molina,^a Taylor Slade,^a Minerva So,^a Karan Thakur,^b Angelica Urquidez Negrete,^a Sage Wackett,^a Sarah Weiss,^a Liam McCarthy,^{b,d} Keith Wheaton,^{b,e} Adam D. Rudner,^{b,e} John P. McCutcheon,^{a,c}  Susanne P. Pfeifer^{a,c,f}

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

^bTranslational and Molecular Medicine Program, University of Ottawa, Ottawa, Ontario, Canada

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

^dDepartment of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada

^eDepartment of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, Canada

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

John P. McCutcheon and Susanne P. Pfeifer jointly supervised the project. Author order was determined alphabetically.

ABSTRACT We characterized the complete genome sequence of *Siphoviridae* bacteriophage Erla, an obligatory lytic subcluster EA1 bacteriophage infecting *Microbacterium foliorum* NRRL B-24224, with a capsid width of 65 nm and a tail length of 112 nm. The 41.5-kb genome, encompassing 62 predicted protein-coding genes, is highly similar (99.52% identity) to that of bacteriophage Calix.

The discovery of antibiotic agents in 1928 by Alexander Fleming has completely changed the way clinicians treat bacterial infections. However, with new therapeutic challenges resulting from antibiotic resistance, scientists have now turned to bacteriophages as promising clinical alternatives to antibiotics (1). Aiding this endeavor is the vast, though largely uncharacterized, diversity of bacteriophages across the globe (current estimates range in the order of 10³¹ [2]).

Here, we report the complete genome sequence of bacteriophage Erla. Erla was obtained from a soil sample collected at 16°C from a garden in Ottawa, Ontario, Canada (45.428778 N, 75.677246 W). Following the standard procedures outlined in the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Sciences (SEA-PHAGES) Discovery Guide (<https://seaphagesphagediscoveryguide.helpdocsonline.com/home>), the sample was mixed in an equal volume of peptone-yeast-calcium agar (PYCa) medium and shaken (250 rpm) at 30°C for 2 h. Next, the sample was filtered (pore size, 0.22 μm), mixed with 500 μl host cells (*Microbacterium foliorum* NRRL B-24224), and grown for 72 h with shaking (220 rpm). A 1-ml sample was filtered (pore size, 0.22 μm), and 100-μl samples of 10-fold serial dilutions were mixed with 300 μl host cells (*Microbacterium foliorum* NRRL B-24224) and 3.5 ml PYCa top agar. This mixture was incubated at 30°C for 2 to 3 days until plaques formed. Erla was purified by repeated rounds of plaque picking into phage buffer, serial dilution, and plaque assays of 100 μl of diluted phage, as described above. Amplification was performed by flooding “webbed” plates with phage buffer overnight at 4°C. Erla forms medium-sized plaques with a faint bullseye in the middle (Fig. 1a) and exhibits a *Siphoviridae* morphology with an icosahedral capsid (diameter, 65 nm) enclosing the double-stranded DNA, attached to a flexible, noncontractile tail (length, 112 nm; Fig. 1b).

Citation Milhaven M, Hastings E, Brister D, Cevallos L, Chilukuri S, Diyya A, Garg A, Hernandez R, Kelly D, Lazo K, Le J, Maag G, Marfatia PD, Mehta R, Nejad A, Porche J, Queiroz A, Sackett D, Santos Molina P, Slade T, So M, Thakur K, Urquidez Negrete A, Wackett S, Weiss S, McCarthy L, Wheaton K, Rudner AD, McCutcheon JP, Pfeifer SP. 2021. Complete genome sequence of *Microbacterium* bacteriophage Erla. *Microbiol Resour Annu* 10:e01354-20. <https://doi.org/10.1128/MRA.01354-20>.

Editor John J. Dennehy, Queens College

Copyright © 2021 Milhaven 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.

Received 10 December 2020

Accepted 6 January 2021

Published 21 January 2021

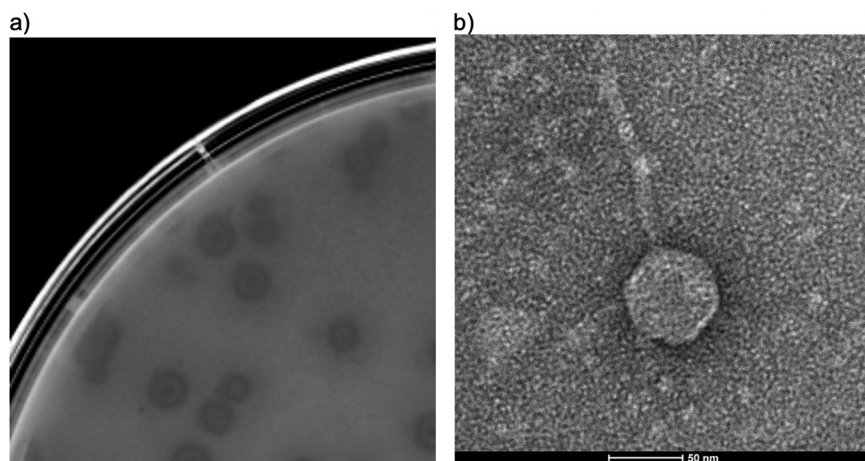


FIG 1 Characterization of the *Microbacterium* bacteriophage Erla. (a) Erla forms medium-sized plaques with a faint bullseye in the middle. (b) Transmission electron microscopy image showing Erla's *Siphoviridae* morphology with an icosahedral capsid (diameter, 65 nm) attached to a 112-nm-long tail.

In order to sequence Erla's genome, DNA was extracted from the phage lysate using a phenol-chloroform assay, followed by ethanol precipitation. Next, a sequencing library was prepared using the NEB Ultra II FS kit and sequenced on an Illumina MiSeq instrument (150-bp single-end reads) to $>3,300\times$ coverage (976,546 reads). Following Russell (3), the raw reads were assembled using Newbler v.2.9 (4), resulting in a single linear contig 41,538 bp in size with a GC content of 63.4%, similar to the 68.7% GC content of its host, *Microbacterium foliorum* NRRL B-24224. The assembly was checked for completeness, accuracy, and genome termini using Consed v.29.0 (5). As neither read start buildups nor substantial variations in coverage were observed, Erla's genome is most likely circularly permuted.

Following the SEA-PHAGES Bioinformatics Guide (<https://seaphagesbioinformatics.helpdocsonline.com/home>), Erla's genome was annotated using DNAMaster v.5.23.5 (<http://cobamide2.bio.pitt.edu>). Sixty-two putative genes were identified using GLIMMER v.3.02 (6), GeneMark v.3.25 and v.4.28 (7), and Starterator v.381 (<https://seaphages.org/software>), corresponding to a gene density of 1.49 genes/kb. ARAGORN v.1.1 (included in DNAMaster) and v.1.2.38 (8) and tRNAscan-SE v.2.0 (9) were used to search for tRNAs and transfer-messenger RNA (tmRNA), but none were found. Functional assignments were made using BLASTp v.2.10.1 (10) and HHpred (11), leading to a putative function for 26 out of the 62 genes. TMHMM v.2.0 (12) and SOSUI v.1.11 (13) were used to gather further information on proteins of no known function, leading to the identification of two additional transmembrane proteins. Default parameters were used for all software.

Phamerator v.381 (14) was used to determine synteny among Erla and other bacteriophages previously sequenced as part of the SEA-PHAGES program, which supported the above functional assignments. Following the classification guidelines established by Hatfull et al. (15), Erla was identified as a *Microbacterium* subcluster EA1 bacteriophage. Multiple sequence alignments using Kalign v.1.04 (16) and BLASTn v.2.10.1 (11) indicated that Erla is most closely related to Calix (percent identity, 99.52%; GenBank accession number [MN234163.1](https://genbank.ncbi.nlm.nih.gov/GenBank/record/MN234163.1)), Gelo (percent identity, 99.51%; [MG962367.1](https://genbank.ncbi.nlm.nih.gov/GenBank/record/MG962367.1)), and Etta (percent identity, 99.51%; [MK977697.1](https://genbank.ncbi.nlm.nih.gov/GenBank/record/MK977697.1)).

Data availability. Whole-genome sequencing data are available through NCBI's Sequence Read Archive (BioProject accession number [PRJNA488469](https://bioinformatics.ncbi.nlm.nih.gov/bioproject/PRJNA488469); run accession number [SRR13108336](https://bioinformatics.ncbi.nlm.nih.gov/sra/record/SRR13108336)). The annotated genome assembly has been deposited in GenBank under accession number [MW291026](https://genbank.ncbi.nlm.nih.gov/GenBank/record/MW291026).

ACKNOWLEDGMENTS

Erla was sequenced as part of the Howard Hughes Medical Institute (HHMI) SEA-PHAGES undergraduate research program. Computations were partially performed at Arizona State University's High Performance Computing facility.

We are grateful to Suhail Ghafoor for IT support, Billy Biederman, Graham Hatfull, Deborah Jacobs-Sera, Welkin Pope, Daniel Russell, and Vic Sivanathan for library preparation, sequencing, and assembly, as well as providing faculty training for the SEA-PHAGES program and guidance with our genome annotation.

This study was supported by Arizona State University's School of Life Sciences.

REFERENCES

- Romero-Calle D, Guimarães Benevides R, Góes-Neto A, Billington C. 2019. Bacteriophages as alternatives to antibiotics in clinical care. *Antibiotics (Basel)* 8:138. <https://doi.org/10.3390/antibiotics8030138>.
- 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).
- 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.
- 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, 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, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380. <https://doi.org/10.1038/nature03959>.
- 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>.
- 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>.
- 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>.
- 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>.
- Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 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>.
- 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>.
- 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>.
- Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko C-C, Weber RJ, Patel MC, Germane KL, Edgar RH, Hoyte NN, Bowman CA, Tantoco AT, Paladin EC, Myers MS, Smith AL, Grace MS, Pham TT, O'Brien MB, Vogelsberger AM, Hryckowian AJ, Wynalek JL, Donis-Keller H, Bogel MW, Peebles CL, Cresawn SG, Hendrix RW. 2010. Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J Mol Biol* 397:119–143. <https://doi.org/10.1016/j.jmb.2010.01.011>.
- Lassmann T, Sonnhammer ELL. 2005. Kalign—an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics* 6:298. <https://doi.org/10.1186/1471-2105-6-298>.