



Complete Genome Sequence of the *Microbacterium* Bacteriophage Chako

Mark Milhaven,^{a,b} Lindsey Cai,^a Sanjana Cherian,^a Kamalei Johnson,^a Kevin Salas Perez,^a Madison Blanco,^a Aman Garg,^a Jackelyn Lobatos,^a Corinne Mitra,^a Maria Strasser,^a Robert Harms,^c  Susanne P. Pfeifer^{a,b}

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

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

^cDepartment of Biology, St. Louis Community College, St. Louis, Missouri, USA

ABSTRACT We characterized the complete genome sequence of Chako, an obligate lytic bacteriophage with siphovirus morphology from subcluster EA1 that infects *Microbacterium foliorum* NRRL B-24224. Its 41.6-kb genome contains 62 putative protein-coding genes and is highly similar to that of bacteriophage HanSolo (99.26% nucleotide identity).

Bacteriophages are a group of viruses that infect bacteria. Because they are increasingly being used to combat antibiotic-resistant bacterial infections (1, 2), characterizing bacteriophage diversity has become an important endeavor.

Here, we report the whole-genome sequence of Chako, a bacteriophage that was isolated from a soil sample that had been collected from a flower bed on the St. Louis Community College campus (38.56799N, 90.42246W) on 5 February 2021. The sample was taken 3 to 5 in. below the surface of the flower bed, which was covered with mulch. Following the protocols outlined in the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) *Phage Discovery Guide* (3), the soil sample was suspended in peptone-yeast extract-calcium (PYCa) liquid medium and shaken at 250 rpm for 1 h at 30°C, the suspension was centrifuged for 10 min at 2,000 × *g*, and the supernatant was filtered (0.22-μm pore size). A plaque assay was then performed by plating an aliquot of the filtrate on PYCa liquid medium with 250 μL of host bacteria (*Microbacterium foliorum* NRRL B-24224), which yielded bacteriophage Chako after 1 to 2 days at 30°C. Chako, which forms small plaques with a clear center surrounded by a turbid border (Fig. 1a), was purified through three rounds of plating. Negative-staining electron microscopy demonstrated that Chako exhibits a siphovirus morphology (Fig. 1b).

DNA was extracted from a lysate of bacteriophage Chako using a Promega Wizard DNA cleanup kit. A sequencing library was prepared using the NEBNext Ultra II library preparation kit (v3 reagents) and sequenced using the Illumina MiSeq platform, which resulted in 59,134 single-end 150-bp reads (approximately 197× coverage). Following the procedure described by Russell (4), the Chako genome was assembled using Newbler v2.9 (5), and the accuracy and completeness of the genome were checked using Consed v29.0 (6). The resulting 41,550-bp contig exhibits a GC content of 63.4%.

Following the protocols outlined in the *SEA-PHAGES Bioinformatics Guide* (7), the Chako genome was annotated using DNA Master v5.23.6 (<http://cobamide2.bio.pitt.edu>), GLIMMER v3.02 (8), GeneMark v2.5 (9), and Starterator v485 (<https://seaphages.org/software/#Starterator>), which identified 62 putative genes. Using BLASTp v2.13.0 (10) and HHpred (11) with the Protein Data Bank (PDB)/MMCF70 (SCOPe70 v2.08 [12 August 2022]), Pfam-A v35, and NCBI Conserved Domain (CD) v3.19 databases, putative functions could be assigned to 26 genes, including several structural and assembly proteins (i.e.,

Editor Simon Roux, DOE Joint Genome Institute

Copyright © 2023 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.

The authors declare no conflict of interest.

Received 6 December 2022

Accepted 20 December 2022

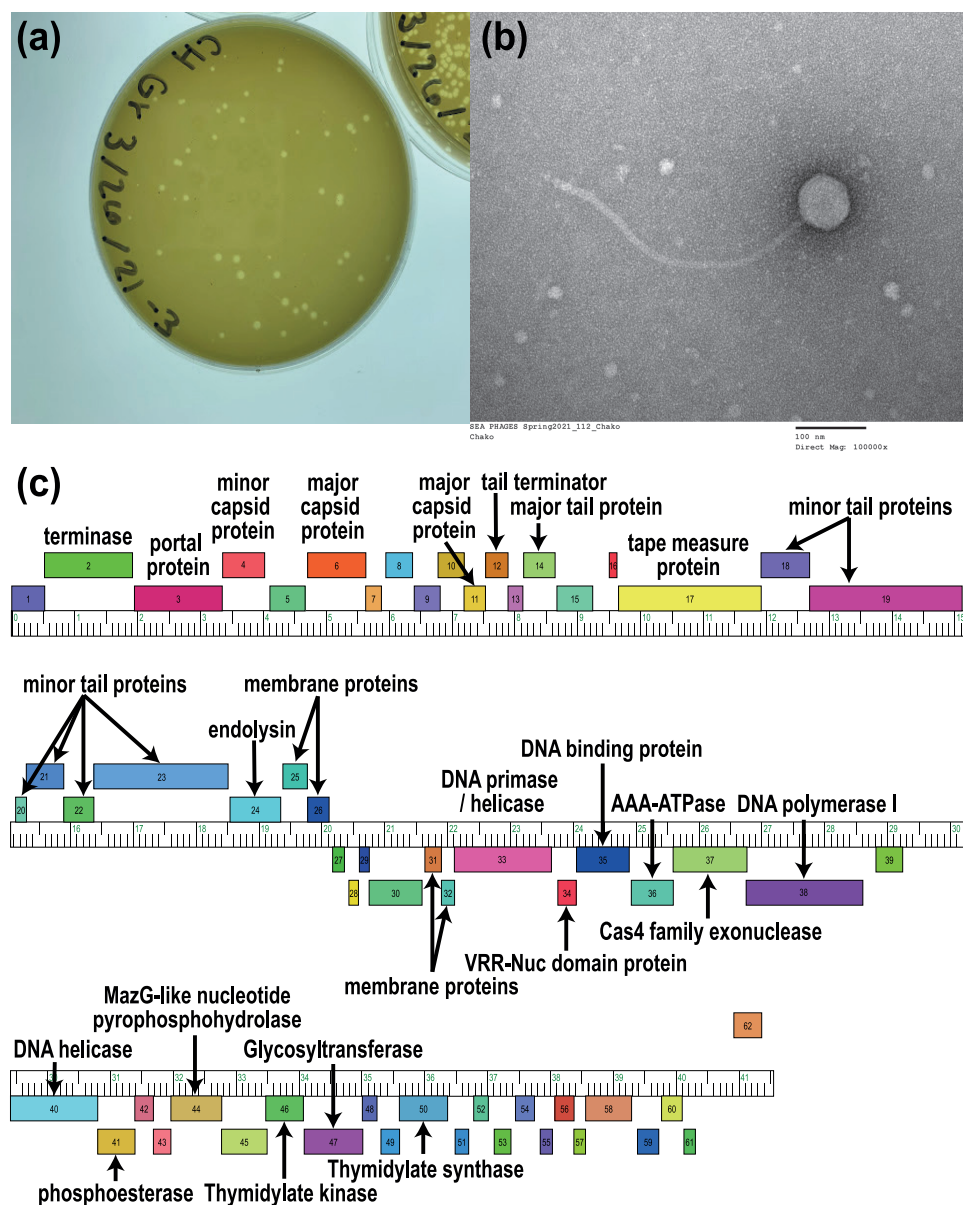


FIG 1 Characterization of the *Microbacterium* bacteriophage Chako. (a) Chako forms small plaques with a clear center surrounded by a turbid border. (b) An electron microscopic image highlights the *Siphoviridae* morphology of Chako, with an icosahedral capsid (diameter, ~65 nm) attached to a noncontractile tail (length, ~420 nm). The sample was viewed at an accelerating voltage of 100 kV with a 1200EX transmission electron microscope (JEOL USA, Peabody, MA) equipped with an 8-megapixel digital camera (Advanced Microscopy Techniques, Woburn, MA) after it was fixed with 1% glutaraldehyde (Ted Pella Inc., Redding, CA) on a freshly glow-discharged Formvar/carbon-coated copper grid for 10 min and stained with 1% aqueous uranyl acetate (Ted Pella Inc.) for 1 min. (c) The Chako genome shows putative protein-coding genes (boxes with numbers representing gene numbers in the genome) on the forward and reverse strands (if available, putative functional assignments are displayed above or below the ruler, respectively).

terminase, portal protein, minor and major capsid proteins, tail terminator, major tail protein, tape measure protein, and six minor tail proteins). Using TMHMM v2.0 (12) and SOSUI v1.11 (13), 4 additional genes could be classified as membrane proteins. ARAGORN v1.2.41 (14) and tRNAscan-SE v2.0 (15) were used to search the genome for tRNAs and transfer-messenger RNAs, but none were found. All software was run with default settings.

Based on gene content similarity of $\geq 35\%$ to bacteriophages in the Actinobacteriophage Database (phagesDB), Chako was assigned to bacteriophage cluster EA1 (16, 17). Consistent with bacteriophages in this cluster, no immunity repressor or integrase functions could be

identified for Chako, suggesting that it is an obligate lytic bacteriophage. Complementary multiple-sequence alignments using Kalign v3.9.01 (18) further demonstrated that Chako is closely related to cluster EA1 bacteriophages HanSolo (GenBank accession number [MK967394](#)) (99.26% identity), Clancy (GenBank accession number [MK967400](#)) (98.90% identity), and Sedgewig (GenBank accession number [MT310851](#)) (98.80% identity).

Data availability. Whole-genome sequencing data are available in the NCBI Sequence Read Archive (SRA) (BioProject accession number [PRJNA488469](#) and SRA accession number [SRR21924933](#)). The annotated genome assembly is available in NCBI GenBank under accession number [OP867023](#).

ACKNOWLEDGMENTS

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

Bacteriophage isolation and electron microscopy were performed at St. Louis Community College (St. Louis, MO), 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 support, Daniel A. Russell and Rebecca A. Garlena for *de novo* assembly, and Billy Biederman, Graham Hatfull, Deborah Jacobs-Sera, and Vic Sivanathan for training and continued support in the SEA-PHAGES program.

REFERENCES

- Kutter E, De Vos D, Gvasalia G, Alavidze Z, Gogokhia L, Kuhl S, Abedon ST. 2010. Phage therapy in clinical practice: treatment of human infections. *Curr Pharm Biotechnol* 11:69–86. <https://doi.org/10.2174/138920110790725401>.
- Kakasis A, Panitsa G. 2019. Bacteriophage therapy as an alternative treatment for human infections: a comprehensive review. *Int J Antimicrob Agents* 53:16–21. <https://doi.org/10.1016/j.ijantimicag.2018.09.004>.
- Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD. <https://seaphagesphagediscoveryguide.helpdocsonline.com/home>.
- Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes, p 109–125. In Clokie MJ, Kropinski AM, Lavigne R (ed), *Bacteriophages: methods and protocols*, vol 3. Springer, New York, NY.
- 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, Jirase 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, Begley RF, Rothberg JM. 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>.
- Pope WH, Jacobs-Sera D, Russell DA, Cresawn SG, Hatfull GF. 2017. SEA-PHAGES bioinformatics guide. Howard Hughes Medical Institute, Chevy Chase, MD. <https://seaphagesbioinformatics.helpdocsonline.com/home>.
- 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>.
- 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>.
- 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>.
- Russell DA, Hatfull G. 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>.
- 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>.