




Complete Genome Sequence of *Streptomyces* Phage Shaeky

Shae Shodrock,^a Tyler Higbee,^a James Clark,^b Isla Hernandez,^b Mei Liu,^b  Ben Burrowes^b

^aDepartment of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, USA

^bCenter for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT Here, we present the genome of siphophage Shaeky, infecting the Gram-positive bacterium *Streptomyces* sp. strain Mg1. Shaeky has very low sequence identity to other phages, with phage phiC31 being the most closely related in the NCBI database. The Shaeky genome is 45,617 bp with 77 protein-coding genes and 16 tRNAs.

Members of *Streptomyces*, the largest genus in the phylum *Actinobacteria*, are characterized as filamentous Gram-positive bacteria. Many of the species are responsible for the decomposition of organic materials in soil (1). Certain species can colonize the respiratory tract and lead to a pulmonary disease known as Cushing syndrome (2).

Bacteriophage Shaeky was isolated by plaque purification, as described previously (3), in February 2019 from a soil sample taken from San Jordan, Utah, using the host strain *Streptomyces* sp. strain Mg1 grown at 30°C on nutrient broth with 10 mM MgCl₂, 8 mM Ca(NO₃)₂, and 0.5% glucose. DNA was purified as described previously (4) using a Wizard DNA clean-up kit (Promega) and was prepared as Illumina TruSeq libraries with 300-bp inserts using a Nextera Flex kit. Sequencing was performed on an Illumina MiSeq system with paired-end 350-bp reads using 500-cycle v2 chemistry. The sequencing reads were quality controlled using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) and then manually trimmed with FastX v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html). The final genome was closed using PCR and Sanger sequencing with forward and reverse primers CGTTGAAGCCGCGATTG and TTCTGCCGAGTTCGTG, respectively. The closed genome was found to have the size of 45,672 bp, with 384.6× contig sequencing coverage, using SPAdes v3.5.0; there were 415,598 total sequencing reads (5). Structural annotations were performed using GLIMMER v3 (6) and MetaGeneAnnotator v1.0 (7). tRNAs were discovered using ARAGORN v2.36 (8). Functionalities of genes were predicted using InterProScan v5.33 (9), TMHMM v2.0 (10), HHPred (11), and BLAST v2.9.0 (12) with the NCBI nonredundant, Swiss-Prot, and TrEMBL databases (13), with default settings (accessed 21 April 2020). The annotation tools described are available on the Galaxy server (<https://cpt.tamu.edu/galaxy-pub>) (14–16).

Genomic analysis predicts phage Shaeky to be a siphophage with a 45,617-bp double-stranded DNA (dsDNA) genome, 77 protein-coding genes, and 16 tRNAs. There were no predicted genome termini.

Of the 77 hypothetical protein-coding genes, 30 had predicted functions while 47 were novel proteins. The 16 tRNAs are located in one sequential stretch that spans ~1.5 kb of the genome. No introns were found. BLASTn analysis shows that Phage Shaeky has only 16.9% nucleotide identity to phage phiC31 (GenBank accession number [AJ006589.3](https://www.ncbi.nlm.nih.gov/nuccore/AJ006589.3)), a *Streptomyces* temperate phage to which Shaeky is most closely related, and shares 25 unique proteins with *Streptomyces* phage Attoomi (GenBank accession number [NC_047905.1](https://www.ncbi.nlm.nih.gov/nuccore/NC_047905.1)). Therefore, Shaeky appears to be a novel siphovirus. All genes in the lysis cassette were identified except for spannins, which are expected for siphophages with Gram-positive hosts. Three predicted HNH endonuclease genes were found, one being

Citation Shodrock S, Higbee T, Clark J, Hernandez I, Liu M, Burrowes B. 2021. Complete genome sequence of *Streptomyces* phage Shaeky. *Microbiol Resour Announc* 10:e01338-20. <https://doi.org/10.1128/MRA.01338-20>.

Editor John J. Dennehy, Queens College

Copyright © 2021 Shodrock 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 Ben Burrowes, benburrowes@tamu.edu.

Received 24 November 2020

Accepted 10 December 2020

Published 7 January 2021

downstream from the span of tRNAs and the other two located toward the beginning of the genome. There is an ~7-kb span in the genome in which all protein-coding genes appear to be novel.

Data availability. The genome of Shaeky was deposited in GenBank with accession number [MT701595.1](https://doi.org/10.12688/f1000research.9534.1). The associated BioProject, SRA, and BioSample accession numbers are [PRJNA222858](https://doi.org/10.1093/nar/27.23.4636), [SRR11558340](https://doi.org/10.1093/nar/27.23.4636), and [SAMN14609632](https://doi.org/10.1093/nar/27.23.4636), respectively.

ACKNOWLEDGMENTS

This work was funded by the National Science Foundation (awards EF-0949351 and DBI-1565146), with additional support from the Center for Phage Technology (CPT).

We thank Paul Straight (Texas A&M University) for providing *Streptomyces* sp. Mg1.

This announcement was prepared in partial fulfillment of the requirements for BICH464 Phage Genomics, an undergraduate course at Texas A&M University.

REFERENCES

1. Chater KF. 2016. Recent advances in understanding *Streptomyces*. *F1000Res* 5:2795. <https://doi.org/10.12688/f1000research.9534.1>.
2. Ataiekhorsangani M, Jafaripozve N, Zaerin O. 2014. *Streptomyces* infection in Cushing syndrome: a case report and literature review. *Adv Biomed Res* 3:26. <https://doi.org/10.4103/2277-9175.124672>.
3. van Charante F, Holtappels D, Blasdel B, Burrowes B. 2019. Isolation of bacteriophages. In Harper D, Abedon S, Burrowes B, McConville M (ed), *Bacteriophages*. Springer, Cham, Switzerland.
4. Summer EJ. 2009. Preparation of a phage DNA fragment library for whole genome shotgun sequencing. *Methods Mol Biol* 502:27–46. https://doi.org/10.1007/978-1-60327-565-1_4.
5. 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>.
6. 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>.
7. Noguchi H, Taniguchi T, Itoh T. 2008. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Res* 15:387–396. <https://doi.org/10.1093/dnares/dsn027>.
8. 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>.
9. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
10. 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>.
11. Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI Bioinformatics Toolkit with a new HHpred server at its core. *J Mol Biol* 430:2237–2243. <https://doi.org/10.1016/j.jmb.2017.12.007>.
12. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
13. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 46:2699. <https://doi.org/10.1093/nar/gky092>.
14. Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A. 2020. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Res* 48:W395–W402. <https://doi.org/10.1093/nar/gkaa434>.
15. Ramsey J, Rasche H, Maughmer C, Criscione A, Mijalis E, Liu M, Hu JC, Young R, Gill JJ. 2020. Galaxy and Apollo as a biologist-friendly interface for high-quality cooperative phage genome annotation. *PLoS Comput Biol* 16:e1008214. <https://doi.org/10.1371/journal.pcbi.1008214>.
16. Dunn NA, Unni DR, Diesh C, Munoz-Torres M, Harris NL, Yao E, Rasche H, Holmes IH, Elsik CG, Lewis SE. 2019. Apollo: democratizing genome annotation. *PLoS Comput Biol* 15:e1006790. <https://doi.org/10.1371/journal.pcbi.1006790>.