AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Streptomyces* Siphophage Sycamore

Xing-Han Zhang,[a] Alejandra Marquez,[a] James Clark,[b] Isla Hernandez,[b] Madeline Rivera,[a] Mei Liu,[b] Ben Burrowes[b]

[a]Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, USA
[b]Center for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT *Streptomyces* sp. strain Mg1 is a competitive soil-dwelling bacterium that secretes antibiotics that inhibit growth of *Bacillus subtilis*. Here, we present the genome sequence of Sycamore, a 44,694-bp *Streptomyces* sp. Mg1 siphophage with 66 predicted protein-coding genes, that is similar to phage genome sequences in the *Lomovskayavirus* genus.

Streptomyces spp. are Gram-positive soil bacteria. Like other streptomycetes, *Streptomyces* sp. strain Mg1 secretes numerous antibiotics that offer the bacterium a growth advantage in the soil environment (1, 2), such as the ability to degrade colonies of *Bacillus subtilis* (2). Here, we describe the isolation and genome annotation of *Streptomyces* sp. Mg1 siphophage Sycamore.

Bacteriophage Sycamore was isolated in February 2019 from an Illinois topsoil sample by plaque purification on *Streptomyces* sp. Mg1 (provided by Paul Straight, Texas A&M University) grown at 30°C on nutrient agar or broth supplemented with 10 mM $MgCl_2$, 8 mM $Ca(NO_3)_2$, and 0.5% glucose using previously reported methods (3). To determine phage morphology, crude Sycamore lysates were stained with 2% (wt/vol) uranyl acetate and viewed via transmission electron microscopy (data not shown) at the Texas A&M Microscopy and Imaging Center (4). DNA was purified as previously described (3) using DNA Wizard DNA clean-up kits and then prepared as Illumina libraries using a Nextera Flex kit to be sequenced on an Illumina MiSeq instrument with paired-end 300-bp reads using V2 500-cycle chemistry. Sequence reads were quality controlled with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) and manually trimmed using FastX 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html). Using SPAdes v3.5.0, a single contig at 275.4-fold coverage was assembled from 686,406 total sequence reads (5). The contig was PCR amplified off the ends (forward primer, 5′-GTAGTGACCACACCCTAGGTAA-3′; reverse primer, 5′-GTATG AGTCGCTGGTCAACAG-3′), and the product was Sanger sequenced to verify sequence closure. Protein-coding genes were predicted with GLIMMER v3 and MetaGeneAnnotator v1.0, tRNAs with ARAGORN v2.36, and rho-independent termination sites with TransTermHP v2.09 (6–9). Functional gene predictions relied on InterProScan v5.33, TMHMM v2.0, and BLAST v2.9.0 (with a 0.001 maximum expectation value cutoff) against the following databases: NCBI nonredundant, UniProtKB Swiss-Prot, and TrEMBL (10–13) (accessed 23 April 2020). Structural predictions were performed with the HHSuite v3.0 tool HHpred (14). The genome-wide DNA sequence similarity to other phages was calculated using progressiveMauve v2.4 (15). Excluding HHpred, all tools were accessed at the Center for Phage Technology Galaxy interface and run with default parameters, and annotation was performed in Web Apollo (hosted online at https://cpt.tamu.edu/galaxy-pub/) (16–18).

Sycamore has a genome size of 44,694 bp with a G+C content of 63%, which is much lower than the characteristically high G+C content observed in *Streptomyces*

Citation Zhang X-H, Marquez A, Clark J, Hernandez I, Rivera M, Liu M, Burrowes B. 2021. Complete genome sequence of *Streptomyces* siphophage Sycamore. Microbiol Resour Announc 10:e01343-20. https://doi.org/10.1128/MRA.01343-20.

Editor Kenneth M. Stedman, Portland State University

Address correspondence to Ben Burrowes, benburrowes@tamu.edu.

Received 24 November 2020
Accepted 10 December 2020
Published 7 January 2021

species (19). Our analysis predicted 68 protein-coding genes, of which 37 were assigned putative functions, and 7 tRNA genes, yielding an overall 90% coding density. A BLASTp search revealed that Sycamore shares the greatest amino acid identity with phages of the *Lomovskayavirus* genus (taxonomy identification number [taxid] 308912), of which *Streptomyces* phage Attoomi (GenBank accession number NC_047905.1) shared the most protein-coding genes, with 25 similar unique proteins. Interestingly, the predicted lysis cassette of Sycamore lacked a predicted holin, and the putative endolysin *N*-acetylmuramidase and two-component spanin were detected approximately 20 kb apart. Moreover, the sequence of a likely tape measure frameshift varied from the 5′-CGGGGGCG-3′ slippery sequence of phage Mu by a single G/A point mutation at the fourth nucleotide (20). No introns were detected.

Data availability. The genome sequence of Sycamore was deposited in GenBank with accession number MT701593.1. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR11558337, and SAMN14609635, respectively.

ACKNOWLEDGMENTS

Funding was provided by the National Science Foundation (awards EF-0949351 and DBI-1565146) and from the Center for Phage Technology (CPT).

Paul Straight, Texas A&M University, provided *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. Hoefler BC, Konganti K, Straight PD. 2013. De novo assembly of the Streptomyces sp. strain Mg1 genome using PacBio single-molecule sequencing. Genome Announc 1:e00535-13. https://doi.org/10.1128/genomeA.00535-13.
2. Barger SR, Hoefler BC, Cubillos-Ruiz A, Russell WK, Russell DH, Straight PD. 2012. Imaging secondary metabolism of Streptomyces sp. Mg1 during cellular lysis and colony degradation of competing Bacillus subtilis. Antonie Van Leeuwenhoek 102:435–445. https://doi.org/10.1007/s10482-012-9769-0.
3. 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.
4. Valentine RC, Shapiro BM, Stadtman ER. 1968. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from Escherichia coli. Biochemistry 7:2143–2152. https://doi.org/10.1021/bi00846a017.
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. Kingsford CL, Ayanbule K, Salzberg SL. 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. Genome Biol 8:R22. https://doi.org/10.1186/gb-2007-8-2-r22.
10. Jones P, Binns D, Chang H-Y, 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.
11. 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.
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. 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.
15. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. https://doi.org/10.1371/journal.pone.0011147.
16. 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.
17. 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.
18. 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.
19. Ohnishi Y, Ishikawa J, Hara H, Suzuki H, Ikenoya M, Ikeda H, Yamashita A, Hattori M, Horinouchi S. 2008. Genome sequence of the streptomycin-producing microorganism Streptomyces griseus IFO 13350. J Bacteriol 190:4050–4060. https://doi.org/10.1128/JB.00204-08.
20. Xu J, Hendrix RW, Duda RL. 2004. Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol Cell 16:11–21. https://doi.org/10.1016/j.molcel.2004.09.006.