Complete Genome Sequence of *Streptomyces* Phage Shady

Carlos Ortega,ᵃ Isla Hernandez,ᵇ James Clark,ᵇ Mei Liu,ᵇ 🔴 Ben Burrowesᵇ

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 *Streptomyces* spp. are saprophytic soil bacteria that produce secondary metabolites with therapeutic potential. This announcement describes the isolation and genome annotation of *Streptomyces* sp. strain Mg1 siphophage Shady. Learning more about Shady's novel 45-kb genome, containing 76 predicted protein-coding genes, could be industrially advantageous when using streptomycetes for their products.

*S*treptomyces spp. are common, soilborne Gram-positive bacteria lacking motility organelles. Many streptomycetes, such as *Streptomyces* sp. strain Mg1, produce secondary metabolites that degrade competitor colonies (1, 2). These metabolites account for over 70% of naturally available antibiotics. Studying bacteriophages that can spoil *Streptomyces* cultures could be beneficial to the drug industry (3–5). This report details the isolation and genome annotation of Shady, a *Streptomyces* sp. Mg1 siphophage.

Shady was isolated from soil collected in Houston, Texas, in August 2019 and propagated on *Streptomyces* sp. Mg1 (provided by Paul Straight, Texas A&M University) by plaque purification in nutrient broth or agar with 10 mM $MgCl_2$, 8 mM $Ca(NO_3)_2$, and 0.5% glucose (6). Samples were negatively stained with 2% (wt/vol) uranyl acetate and imaged at the Texas A&M Microscopy and Imaging Center using transmission electron microscopy (data not shown). Shady DNA was purified using a modified Wizard DNA clean-up kit (Promega) protocol as described elsewhere (7). iSeq libraries were prepared with 300-bp inserts using a TruSeq Nano kit and sequenced using V2 500-cycle chemistry on an Illumina MiSeq instrument with paired-end 250-bp reads. FastQC was used to visualize and quality control the 325,742 reads (www.bioinformatics.babraham.ac.uk/projects/fastqc) which were then trimmed with FastX Toolkit 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html), and one contig was assembled at 107.3-fold coverage using SPAdes v3.5.0 (8). The genome was closed using PCR and Sanger sequencing, (forward primer, 5′-GAGGAATCCATCTAGCGAAGTC-3′; reverse primer, 5′-ATAATGCGGATCTGTTCGGG-3′). PhageTerm was used to predict the genomic termini (9). MetaGeneAnnotator v1.0 (10) and GLIMMER v3 (11) were used to predict protein-coding genes, and ARAGORN v2.36 (12) was used to detect tRNA genes. To predict gene function, conserved domain searches with InterProScan v5.33 (13), transmembrane domain searches with TMHMM v2.0 (14), and similarity searches with BLAST v2.9.0 (15) versus the NCBI nonredundant (nr), Swiss-Prot, and TrEMBL databases (16) were utilized at the default settings (accessed 24 March 2020). Similarity calculations of the genome-wide DNA sequence were carried out using progressiveMauve v2.4 (17). Tools were accessed via the Center for Phage Technology Web Galaxy interface online at https://cpt.tamu.edu/galaxy-pub, using Web Apollo for annotations (18–20). Default parameters were used for all software unless otherwise specified.

The 45,128-bp double-stranded DNA (dsDNA) genome of Shady has a GC content of 62.3%, which is significantly less than the 72.2% of its host (1). Genomic analysis predicted 76 protein-coding and 8 tRNA genes, yielding a coding density of 89.3%. PhageTerm could not detect genomic termini.

Of the 76 identified protein-coding genes, 33 were assigned putative functions.

Various putative tail protein genes were found with high similarity to *Mycobacterium* phages Bxb1 and L5. They include two tape measure protein chaperone genes embedded in the same exon with a slippery sequence causing a frameshift resulting in two different products. Putative DNA primase and DNA polymerase I genes were identified, indicating that Shady might form at least some of its own DNA replication machinery. Furthermore, a putative integrase gene very closely related to lambdoid phage phiC31 was found, suggesting that Shady is a temperate phage (4). Transmission electron microscopy showed the virion to be a siphovirus, and comparative genomics revealed Shady to have no close relatives. Shady is therefore a novel *Streptomyces* siphophage.

Data availability. The genome sequence of Shady is available in GenBank (accession number MT701596.1). The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR11558351, and SAMN14609630, respectively.

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. Torres-Barceló C. 2018. The disparate effects of bacteriophages on antibiotic-resistant bacteria. Emerg Microbes Infect 7:168. https://doi.org/10.1038/s41426-018-0169-z.
4. Hendrix RW, Hatfull GF, Ford ME, Smith MCM, Burns RN. 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc Natl Acad Sci U S A 96:2192–2197. https://doi.org/10.1073/pnas.96.5.2192.
5. Al-Ansari M, Alkubaisi N, Vijayaragavan P, Murugan K. 2019. Antimicrobial potential of Streptomyces sp. to the Gram positive and Gram negative pathogens. J Infect Public Health 12:861–866. https://doi.org/10.1016/j.jiph.2019.05.016.
6. Adams MH. 1956. Bacteriophages. Interscience Publishers, Inc., New York, NY.
7. 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.
8. 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.
9. Garneau JR, Depardieu F, Fortier L-C, Bikard D, Monot M. 2017. PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. Sci Rep 7:8292. https://doi.org/10.1038/s41598-017-07910-5.
10. 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.
11. 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.
12. 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.
13. 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 S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240. https://doi.org/10.1093/bioinformatics/btu031.
14. Krogh A, Larsson B, 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.
15. 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.
16. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
17. 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.
18. 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.
19. 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.
20. 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.