



# Genome Sequence of a Cluster CR2 *Gordonia terrae* Phage, StarStruck

Wyatt Cannell,<sup>a</sup> Eleanor Carrolton,<sup>a,b</sup> Julia Coombs,<sup>a</sup> Addison Gambol,<sup>a</sup> McHenna Martin,<sup>a,b</sup>  Melody Neely,<sup>a</sup>  Sally Molloy<sup>a,b</sup>

<sup>a</sup>Molecular and Biomedical Sciences, University of Maine, Orono, Maine, USA

<sup>b</sup>The Honors College, University of Maine, Orono, Maine, USA

**ABSTRACT** Bacteriophage StarStruck is a lytic *Siphoviridae* phage that infects *Gordonia terrae* 3612. The 68,128-bp genome of StarStruck has a GC content of 65.4% and contains 92 protein-coding genes, including the gene for a HicA-like toxin. StarStruck was assigned to subcluster CR2 based on >35% shared gene content with other cluster CR genomes in the Actinobacteriophage Database.

Actinobacteriophages, which are viruses that infect bacteria of the phylum *Actinobacteria*, are extremely abundant and diverse (1–4). By characterizing lytic bacteriophages, we increase our understanding of phage diversity and susceptibility, which is relevant to applications such as phage therapy and biological control of foaming bacteria in wastewater treatment plants (5, 6). StarStruck was isolated from a soil sample collected at 27°C in Orono, Maine (44.896455N, 68.66614W), using *Gordonia terrae* 3612 (7). Soil extracts were prepared in peptone-yeast extract-calcium (PYCa) medium and filtered on 0.22- $\mu$ m filters before inoculation with *G. terrae* and incubation at 30°C for 2 days. The extract was diluted and plated onto PYCa agar in soft agar containing *G. terrae*, and plaques were purified by four rounds of plaque assays (7). After incubation for 2 days at 30°C, StarStruck formed 4-mm clear plaques with turbid edges on a lawn of *G. terrae*, but lysogen isolation techniques did not yield stable lysogens (7). The particle morphology of StarStruck was determined by negative staining transmission electron microscopy. StarStruck has a *Siphoviridae* morphology with a 65-nm (standard error [SE],  $\pm 3.2$  nm) icosahedral head and a 297.5-nm (SE,  $\pm 7.5$  nm) flexible, noncontractile tail ( $n = 3$  particles).

DNA was extracted from a high-titer lysate by phenol-chloroform extraction and prepared for sequencing using the NEBNext Ultra II library preparation kit (New England BioLabs, Ipswich, MA) (8). Sequencing on an Illumina MiSeq platform yielded 561,846 single-end 150-bp reads. Newbler v2.9 and Consed v29 (9) were used to assemble raw reads and check for completeness, yielding a 68,128-bp genome with a GC content of 65.4%. Genome ends are defined by single-stranded 10-bp 3' extensions (CGCCGCGTAC). StarStruck shares >35% gene content with members of cluster CR in the Phamerator database Actino\_Draft and was assigned to subcluster CR2 (4, 10, 11).

The genome of StarStruck was autoannotated using GLIMMER v3.02 and GeneMark v2.5 within DNA Master v5.23.6 (<http://cobamide2.bio.pitt.edu>) and PECAAN (<https://blog.kbringsd.org/>) (12, 13). Translational starts were predicted manually based on inclusion of coding potential in GeneMark.hmm and conservation across homologs in BLAST and Starterator (<http://phages.wustl.edu/starterator>) analyses (14). Putative gene functions were predicted using BLAST, TMHMM, and HHPred, and gene maps were prepared using the Phamerator database Actino\_Draft (10, 15, 16). No tRNA genes were identified by ARAGORN v1.2.38 and tRNAscan-SE (17, 18). StarStruck contains 92 protein-coding genes. The left arm of the genome contains forward-transcribed genes (gp1 to gp53) beginning with the small-subunit terminase (gp1) (Fig. 1). The large-subunit terminase and the rest of the structural and assembly genes (gp25 to gp46) continue after an

**Editor** John J. Dennehy, Queens College CUNY

**Copyright** © 2022 Cannell 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 Sally Molloy, [sally.dixon@maine.edu](mailto:sally.dixon@maine.edu).

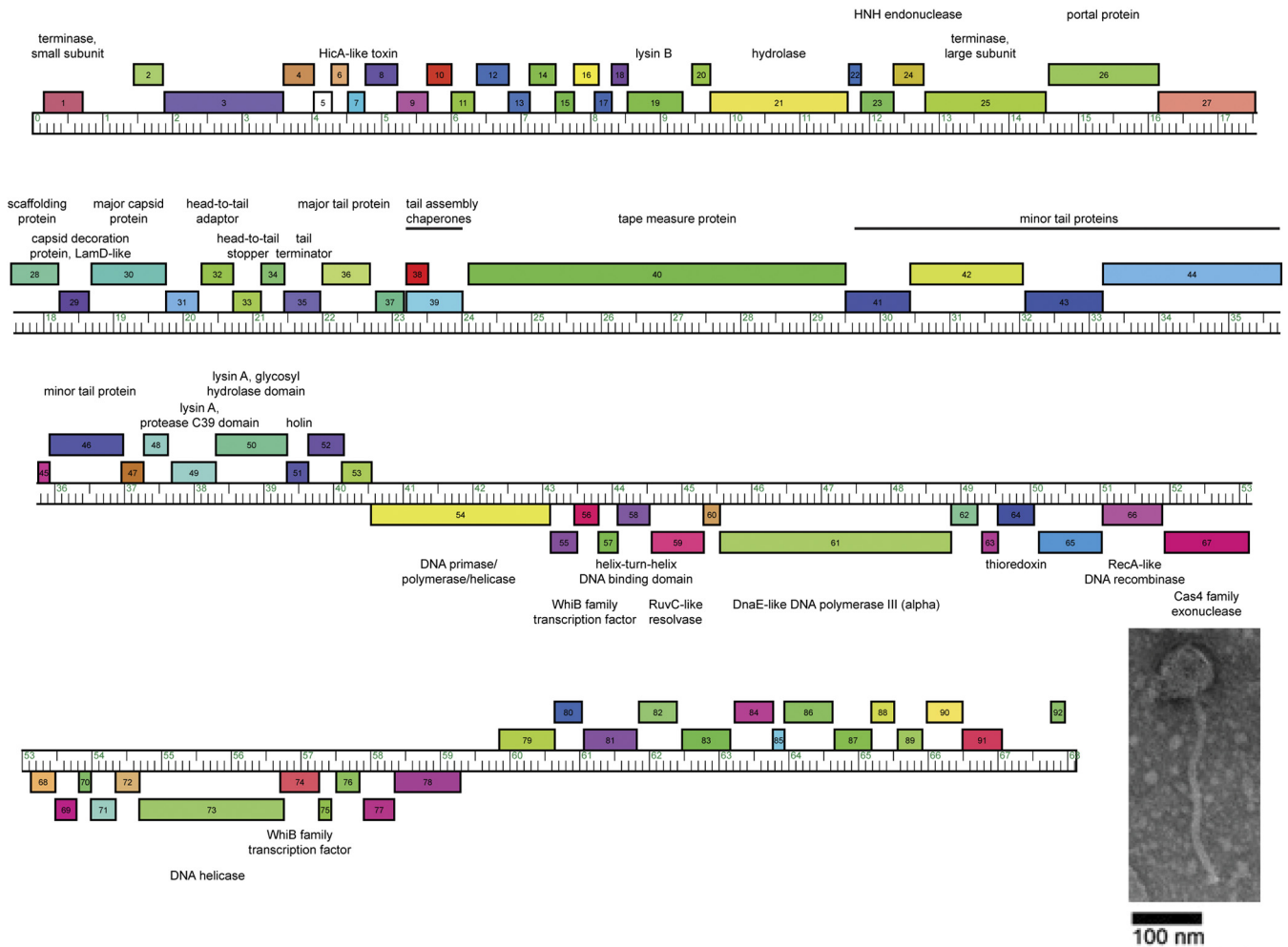
The authors declare no conflict of interest.

**Received** 22 July 2022

**Accepted** 15 August 2022

**Published** 30 August 2022

StarStruck (CR2)



**FIG 1** Genome map of *Gordonia* phage StarStruck. The ruler indicates genome coordinates in units of kilobase pairs. Forward and reverse genes are represented by colored boxes above and below the ruler, respectively. Genes were assigned to a phamily using Phamerator (10) with the Actino\_Draft database, and different phamilies are indicated by different colors. An electron micrograph of StarStruck is shown in the inset, with a scale bar of 100 nm. The average diameter and length of the particle head and tail were determined by measuring the dimensions of three different particles.

~12,000-bp region containing 23 genes. The right arm contains reverse-transcribed genes (gp54 to gp92), including two WhiB family transcription factors (gp56 and gp74), a DnaE-like DNA polymerase (gp61), and a DNA helicase (gp73). StarStruck lacks lysogenic genes such as an integrase and an immunity repressor, indicating that StarStruck is a lytic phage (19).

Like many cluster CR phages, StarStruck encodes a HicA-like toxin (gp7) within the 12,000-bp region separating the small- and large-subunit terminases. Another interesting feature of StarStruck is the location of the lysin B (gp19) within this region, rather than adjacent to the lysin A genes (protease C39 domain [gp49] and glycosyl hydrolase domain [gp50]), which are located downstream of the minor tail proteins.

**Data availability.** StarStruck is available at GenBank with the accession number [ON456333](https://www.ncbi.nlm.nih.gov/nuccore/ON456333) and the Sequence Read jhu7 (SRA) accession number [SRX14816101](https://www.ncbi.nlm.nih.gov/sra/SRX14816101).

**ACKNOWLEDGMENTS**

This research was made possible by the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program of the Howard Hughes Medical Institute. We thank Daniel Russel and Rebecca Garlena for sequencing services and assembly of the StarStruck genome. We are grateful to Geoff Williams at the Brown Bioimaging Facility for providing electron microscopy services. Students are

grateful for the support of teaching assistants Remi Geohagen, Mathew Cox, Allie Conner, Caitlin Wiafe-Kwakye, Abigail McNally, and Jacob Cote.

Research reported in this project was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant P20GM103423.

## REFERENCES

- Hatfull GF. 2010. Mycobacteriophages: genes and genomes. *Annu Rev Microbiol* 64:331–356. <https://doi.org/10.1146/annurev.micro.112408.134233>.
- 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>.
- Jacobs-Sera D, Abad LA, Alvey RM, Anders KR, Aull HG, Bhalla SS, Blumer LS, Bollivar DW, Bonilla JA, Butela KA, Coomans RJ, Cresawn SG, D'Elia T, Diaz A, Divens AM, Edgington NP, Frederick GD, Gainey MD, Garlena RA, Grant KW, Gurney SMR, Hendrickson HL, Hughes LE, Kenna MA, Klyczek KK, Kotturi H, Mavrich TN, McKinney AL, Merkhofer EC, Moberg Parker J, Molloy SD, Monti DL, Pape-Zambito DA, Pollenz RS, Pope WH, Reyna NS, Rinehart CA, Russell DA, Shaffer CD, Sivanathan V, Stoner TH, Stukeley J, Sunnen CN, Tolsma SS, Tsourkas PK, Wallen JR, Ware VC, Warner MH, Washington JM, Westover KM, Whitefleet-Smith JL, Wiersma-Koch HI, Williams DC, Zack KM, Hatfull GF. 2020. Genomic diversity of bacteriophages infecting *Mycobacterium* spp. *PLoS One* 15:e0234636. <https://doi.org/10.1371/journal.pone.0234636>.
- Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, Russell DA, Warner MH, Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEAPHAGES), Hatfull GF. 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>.
- Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, Gilmour KC, Soothill J, Jacobs-Sera D, Schooley RT, Hatfull GF, Spencer H. 2019. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat Med* 25:730–733. <https://doi.org/10.1038/s41591-019-0437-z>.
- Dyson ZA, Tucci J, Seviour RJ, Petrovski S. 2015. Lysis to kill: evaluation of the lytic abilities, and genomics of nine bacteriophages infective for *Gordonia* spp. and their potential use in activated sludge foam biocontrol. *PLoS One* 10:e0134512. <https://doi.org/10.1371/journal.pone.0134512>.
- Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD.
- Sambrook J, Russell DW. 2006. Purification of nucleic acids by extraction with phenol:chloroform. *Cold Spring Harb Protoc* 2006:pdb.prot4455. <https://doi.org/10.1101/pdb.prot4455>.
- Gordon D, Green P. 2013. Conseq: a graphical editor for next-generation sequencing. *Bioinformatics* 29:2936–2937. <https://doi.org/10.1093/bioinformatics/btt515>.
- 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>.
- Russell DA, Hatfull GF. 2017. PhagesDB: the Actinobacteriophage Database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.
- Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* 33:W451–W454. <https://doi.org/10.1093/nar/gki487>.
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673–679. <https://doi.org/10.1093/bioinformatics/btm009>.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
- 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>.
- 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>.
- 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>.
- Mageeney C, Pope WH, Harrison M, Moran D, Cross T, Jacobs-Sera D, Hendrix RW, Dunbar D, Hatfull GF. 2012. Mycobacteriophage Marvin: a new singleton phage with an unusual genome organization. *J Virol* 86:4762–4775. <https://doi.org/10.1128/JVI.00075-12>.