Genome Sequence of Singleton *Gordonia terrae* Bacteriophage Finkle

Seth Ashby,a Gavin Bressette,a Sydney Brown,a Sophie Charles,a Melissa Maginnis,a Sally D. Molloy,a,b Melody N. Neelya

aMolecular and Biomedical Sciences, University of Maine, Orono, Maine, USA
bThe Honors College, University of Maine, Orono, Maine, USA

ABSTRACT Bacteriophage Finkle is a temperate siphovirus isolated from soil on *Gordonia terrae*. The 47,895-bp genome has a GC content of 66.6% and encodes 84 protein-coding genes. The genome is not closely related to sequences in the Actinobacteriophage database, sharing less than 35% gene content, and was classified as a singleton.

Actinobacteriophages (phages) are diverse viruses that infect bacteria of the phylum *Actinobacteria* (1–3). Studying the diversity of phages increases our understanding of phage evolution and is a valuable platform for science education (4). Actinobacteriophage Finkle was isolated from garden soil collected on 01 September 2021 in Brewer, Maine (44.808342 N, 68.729118 W), on the host *Gordonia terrae* 3612. Soil extracts were prepared in peptone-yeast extract-calcium (PYCa) media and filtered on 0.22-μM filters before inoculating with *G. terrae*. After incubating at 30°C for 2 days, the extract was diluted and plated in soft agar containing *G. terrae* onto PYCa agar. Finkle plaques were purified by three rounds of plaque assays resulting in turbid plaques 1 to 3 mm in diameter on a lawn of *G. terrae* after 2 days incubation at 30°C (5). The Finkle lysate was examined by negative-stained transmission electron microscopy revealing a siphovirus morphology with a long, flexible, noncontractile tail 324 $\pm$ 11.5 SE nm in length and an icosahedral head of 62 $\pm$ 6.0 SE nm in diameter (*n* = 2).

DNA was extracted from high-titer lysate by phenol-chloroform extraction (6). DNA was prepared for sequencing using the Ultra II library kit (New England BioLabs [NEB], Ipswich, MA) and sequenced on an Illumina MiSeq platform. This process yielded 19,826 single-end 150-bp reads. Newbler v2.9 and Consed v29 were used to assemble the genome sequence and check for completeness, yielding a 47,895-bp genome with 29-fold coverage and a GC content of 66.6%. Genome ends are defined by single-stranded 14-bp 3′ extensions (CTACCTGCGGGGGA). The genome did not share 35% gene content with sequences in the Phamerator Actino_Draft database and was classified as a singleton, which is a phage without close relatives (3, 7, 8).

Autoannotations of the Finkle genome were performed using GLIMMER v3.02 and GeneMark v2.5 within programs DNA Master v5.23.6 (http://cobamide2.bio.pitt.edu) and PECAAN (https://blog.kbrinsgd.org/) (9, 10). Translational starts were determined manually by identifying starts that include the coding potential predicted in GeneMark.hmm and are conserved across homologs in BLASTp and Starterator (http://phages.wustl.edu/starterator/) analyses (11). Gene functions were predicted using BLASTp, TMHMM, and HHpred (12, 13). No tRNA genes were identified by ARAGORN v1.2.38 and tRNAscan-SE (14, 15). All tools were run with default parameters unless otherwise specified. Comparative genomics was performed using Phamerator in the Actino_Draft database (7). The genome encodes 84 protein-coding genes (Fig. 1). The left arm encodes forward-transcribed structural and assembly genes (gp1 to 25) that belong to gene Phams that is common in mycobacteriophages and *Gordonia* phage clusters (e.g., clusters CZ and P). The right arm encodes forward-transcribed genes
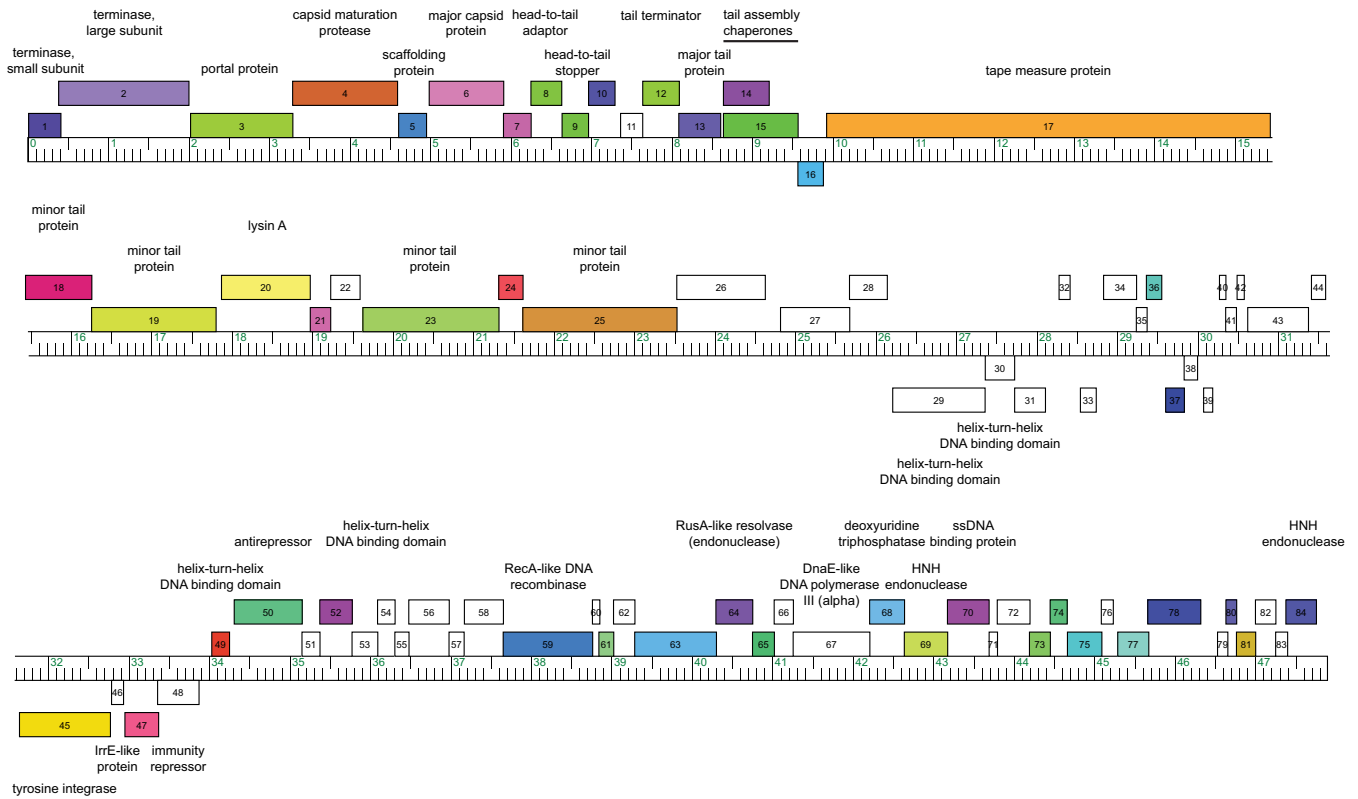
FIG 1 Genome map of *Gordonia* phage Finkle. The genome coordinates are represented by the ruler with kilobase pair units. Forward and reverse genes are represented by colored boxes above and below the ruler, respectively. Genes were assigned to a phamily using Phamerator (7) in the Actino_draft database, and different phamilies are indicated by colors. Genes belonging to Orphams are indicated by white boxes.

(gp49 to 84) including helix-turn-helix DNA binding proteins (gp49 and 54) and a RecA-like DNA recombinase (gp59).

Finkle encodes 37 genes belonging to orphams, which are phams with a single gene member (3). Many of them are located in the center of the genome (23,500 to 34,000 bp) and include both forward- and reverse-transcribed genes. The function of many of these genes is unknown but includes two helix-turn-helix DNA binding domains (gp29 and 30), a tyrosine integrase (gp45), and the immunity repressor (gp48), indicating that Finkle is a temperate phage (16). The immunity repressor is also an orpham; however, homologs were identified within the NCBI database in prophage sequences within bacterial genomes, including *Corynebacterium nuruki* (GenBank accession no. NZ_CP042429.1), *Gordonia iterans* (NZ_CP027433.1), and *Rhodococcus equi* (NZ_LWTY01000001.1).

Data availability. Finkle is available at GenBank with the accession no. ON456347 and at the Sequence Read Archive (SRA) with no. SRX14443501.

REFERENCES

1. 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, Stukey J, Sunnen CN, Tolsma SS, Tsourkas PK, Wallen JR, Ware VC, Warner MH, Washington JM, Westover KM, et al. 2020. Genomic diversity of bacteriophages infecting Microbacterium spp. PLoS One 15:e0234636. https://doi.org/10.1371/journal.pone.0234636.

2. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs WR, Jr, Hendrix RW, Lawrence JG, Hatfull GF, Mycobacterial Genetics Course. 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife 4:e06416. https://doi.org/10.7554/eLife.06416.

3. Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, Russell DA, Warner MH, Genomics S-P, Science E, 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.

4. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, Findley AM, Gissendanner CR, Golebiewska UP, Guild N, Hartzog GA, Grillo WH, Hollowell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenzweig F, Rubin MR, Saha MS, Sandoz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradley KW, Jacobs-Sera D, Pope WH, Russell DA, Cresawn SG, Lopatto D, Bailey CP, Hatfull GF. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. mBio 5:e01051-13. https://doi.org/10.1128/mBio.01051-13.

5. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD.

6. 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.

7. 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.

8. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

15. 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.

16. Hatfull GF. 2010. Mycobacteriophages: genes and genomes. Annu Rev Microbiol 64:331–356. https://doi.org/10.1146/annurev.micro.112408.134233.