



# Genome Sequence of a Cluster DN1 *Gordonia terrae* Phage, Periwinkle

Emma M. Boudreaux,<sup>a,b</sup> Sophie B. Childs,<sup>a,b</sup> Abigail Dichiaro,<sup>a</sup> Amy Hardy,<sup>a,b</sup> Sam Kovacs,<sup>a</sup> Alison F. Kueck,<sup>a,b</sup> Parker M. Landesbergen,<sup>a</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** Periwinkle is a temperate bacteriophage that was isolated on the host *Gordonia terrae* 3612. The genome has a length of 55,657 bp and a GC content of 62.9% and contains 109 protein-coding genes and no tRNA genes. An 8-kb region after the structural protein genes encodes eight membrane proteins, a tyrosine integrase, and an immunity repressor.

**A**ctinobacteriophages are extremely abundant and diverse viruses that infect bacteria within the phylum *Actinobacteria* (1–4). By studying bacteriophages through programs such as the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) Program, we advance our understanding of phage and bacterial diversity, evolution, and virus-host interactions (3–6). Periwinkle was isolated from a composted manure sample collected in Orono, Maine (44.915628N, 68.69072W), using *Gordonia terrae* 3612 (7). Soil extracts were prepared in peptone-yeast extract-calcium (PYCa) medium, filtered on 0.22- $\mu$ m filters, inoculated with *G. terrae*, and incubated at 30°C for 48 h. Dilutions of the enriched extract were plated onto PYCa agar in soft agar containing *G. terrae*, and plaques were purified by five rounds of plaque assays (7). Periwinkle formed 3-mm turbid plaques on a lawn of *G. terrae* (7). Periwinkle has a *Siphoviridae* particle morphology, as determined by negative staining transmission electron microscopy. The particle has a 65-nm (standard error [SE],  $\pm 0.8$  nm) icosahedral head and a 335-nm (SE,  $\pm 8.2$  nm) flexible, noncontractile tail ( $n = 5$ ).

A phenol-chloroform extraction method was used to extract DNA from a high-titer lysate before it was prepared for sequencing using the NEBNext Ultra II library preparation kit (New England BioLabs, Ipswich, MA) (8). Sequencing on an Illumina MiSeq platform yielded 168,288 single-end 150-bp reads. Newbler v2.9 and Consed v29 (9) were used for *de novo* assembly and checks for completeness, yielding a 55,657-bp genome with a GC content of 58.1%. Genome ends are defined by single-stranded 10-bp 3' extensions (CTCGGGGCAT). Periwinkle shares >35% gene content with members of cluster DN in the Phamerator Actino\_Draft database and was assigned to subcluster DN1 (4, 10, 11).

The genome of Periwinkle 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.kbrinsgd.org/>) before manual refining of translational starts based on inclusion of coding potential predicted by GeneMark.hmm and conservation across homologs according to BLAST and Starterator (<http://phages.wustl.edu/starterator>) (12–14). 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). Periwinkle contains 109 protein-coding genes. The left arm of the genome contains mainly forward-transcribed assembly and structural genes (gp1 to gp36) (Fig. 1). All cluster DN phages contain at least one

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

**Copyright** © 2022 Boudreaux 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).

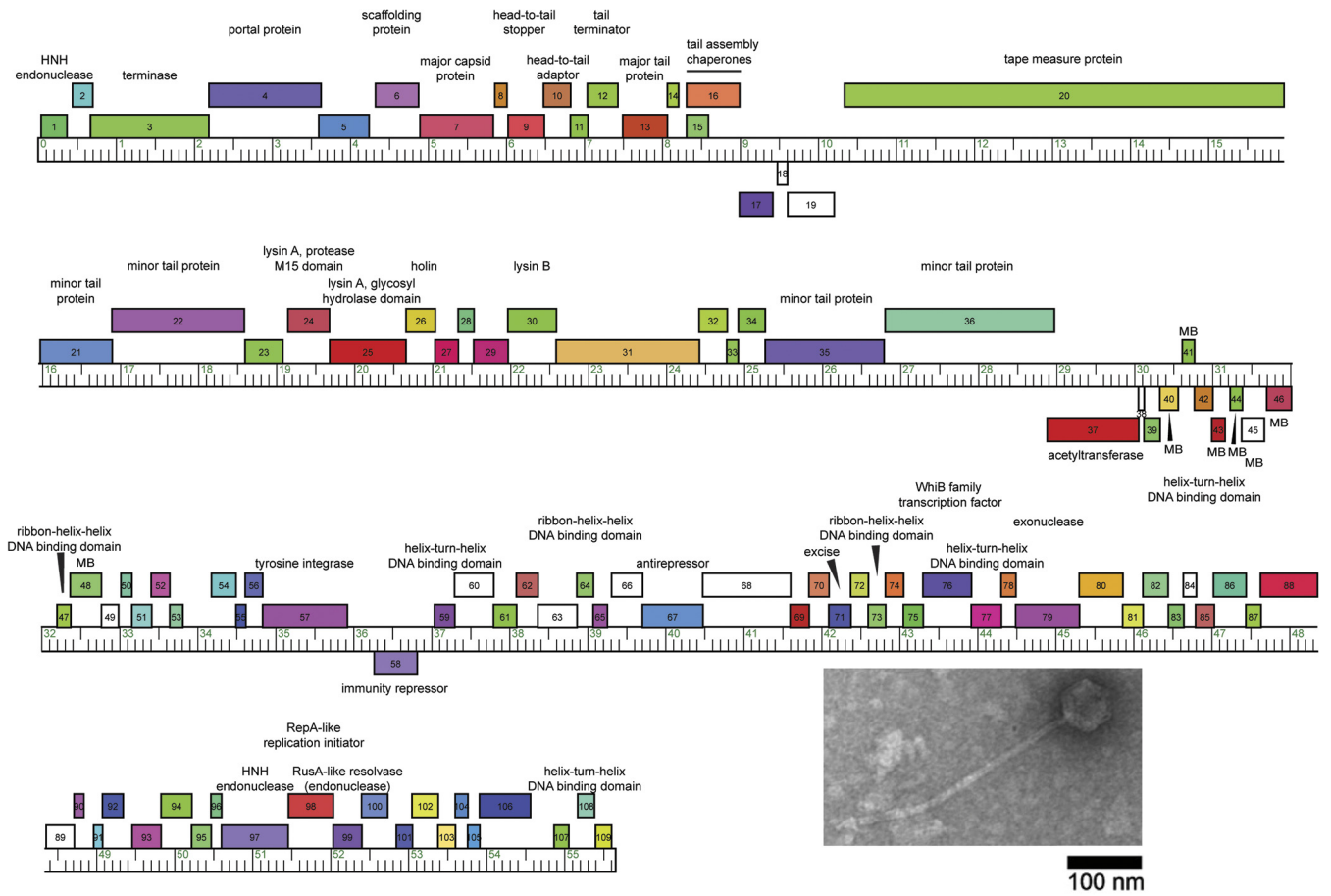
The authors declare no conflict of interest.

**Received** 22 July 2022

**Accepted** 15 August 2022

**Published** 25 August 2022

### Periwinkle (DN1)



**FIG 1** Genome map of *Gordonia* phage Periwinkle. The genome coordinates are represented by the ruler, 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. Gene phamilies with only a single gene member (orphams) are represented by white boxes. Genes with transmembrane domains are labeled MB. An electron micrograph of Periwinkle is shown in the inset, with a scale bar of 100 nm.

reverse-transcribed gene between the tail assembly chaperones and the tape measure protein, and Periwinkle contains three such genes (gp17 to gp19), including two orphams, i.e., gene phamilies with one member in the Actinobacteriophage Database (4, 11). The right arm contains forward-transcribed genes (gp59 to gp109), including five DNA-binding proteins (gp59, gp64, gp73, gp77, and gp108), an antirepressor (gp67), and a WhiB family transcription factor (gp76); gp57 and gp58 encode a tyrosine integrase and an immunity repressor, respectively, indicating that Periwinkle is likely a temperate phage (1).

The integrase and immunity repressor genes are located in an 8-kb region following the minor tail protein genes (gp35 and gp36) that contains forward- and reverse-transcribed genes (gp37 to gp58) that are likely expressed during lysogeny (19). Included are two DNA-binding proteins (gp42 and gp47), an acetyltransferase (gp37), and seven membrane proteins (gp40, gp41, gp44 to gp46, and gp48) that could contribute to superinfection immunity (19).

**Data availability.** Periwinkle is available in GenBank with the accession number [ON456334](https://www.ncbi.nlm.nih.gov/nuclink/ON456334) and the Sequence Read Archive (SRA) accession number [SRR18715698](https://www.ncbi.nlm.nih.gov/sra/SRR18715698).

### ACKNOWLEDGMENTS

This research was made possible by the SEA-PHAGES program at the Howard Hughes Medical Institute. We thank Daniel Russel and Rebecca Garlena for sequencing services in assembling the Periwinkle 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, Stukej 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 *Microbacterium* 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 (SEA-PHAGES), 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>.
- Hatfull GF. 2013. Complete genome sequences of 63 mycobacteriophages. *Genome Announc* 1:e00847-13. <https://doi.org/10.1128/genomeA.00847-13>.
- 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>.
- 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. Consed: 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>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- 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, 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>.
- 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>.
- Dedrick RM, Jacobs-Sera D, Bustamante CAG, Garlena RA, Mavrich TN, Pope WH, Reyes JCC, Russell DA, Adair T, Alvey R, Bonilla JA, Bricker JS, Brown BR, Byrnes D, Cresawn SG, Davis WB, Dickson LA, Edgington NP, Findley AM, Golebiewska U, Grose JH, Hayes CF, Hughes LE, Hutchison KW, Isern S, Johnson AA, Kenna MA, Klyczek KK, Mageeey CM, Michael SF, Molloy SD, Montgomery MT, Neitzel J, Page ST, Pizzorno MC, Poxleitner MK, Rinehart CA, Robinson CJ, Rubin MR, Teyim JN, Vazquez E, Ware VC, Washington J, Hatfull GF. 2017. Prophage-mediated defence against viral attack and viral counter-defence. *Nat Microbiol* 2:16251. <https://doi.org/10.1038/nmicrobiol.2016.251>.