

Genome Sequence of *Gordonia* Phage Emalyn

Welkin H. Pope, Madeline J. Guido, Pragnya Iyengar, Jonathan T. Nigra, Matthew B. Serbin, Naomi S. Kasturiarachi, Catherine A. Pressimone, Johnathon G. Schiebel, Emily C. Furbee, Sarah R. Grubb, Marcie H. Warner, Matthew T. Montgomery, Rebecca A. Garlena, Daniel A. Russell, Deborah Jacobs-Sera, Graham F. Hatfull

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Emalyn is a newly isolated bacteriophage of *Gordonia terrae* 3612 and has a double-stranded DNA genome 43,982 bp long with 67 predicted protein-encoding genes, 32 of which we can assign putative functions. Emalyn has a prolate capsid and has extensive nucleotide similarity with several previously sequenced phages.

Received 9 May 2016 Accepted 21 June 2016 Published 11 August 2016

Citation Pope WH, Guido MJ, Iyengar P, Nigra JT, Serbin MB, Kasturiarachi NS, Pressimone CA, Schiebel JG, Furbee EC, Grubb SR, Warner MH, Montgomery MT, Garlena RA, Russell DA, Jacobs-Sera D, Hatfull GF. 2016. Genome sequence of *Gordonia* phage Emalyn. *Genome Announc* 4(4):e00597-16. doi:10.1128/genomeA.00597-16.

Copyright © 2016 Pope 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 Welkin H. Pope, welkin@pitt.edu.

Gordonia spp. are actinomycetes found in soil and are associated with foaming in activated sludge of wastewater treatment plants (1). Bacteriophages of *Gordonia* hosts have been isolated and seventeen genomes have been sequenced (2–6) but whether the genetic diversity of *Gordonia* phages is as varied as the large collection of mycobacteriophages (7) remains ill-defined. The Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program provides a platform for exploring phages of actinobacterial hosts including phages of *Gordonia* (7, 8).

Phage Emalyn was isolated from a flower bed in Pittsburgh, Pennsylvania using *Gordonia terrae* 3612 as host. Plaques were recovered by direct plating of filtered soil extracts on lawns of *Gordonia*; phages were purified, amplified, and dsDNA was extracted. Emalyn forms relatively small plaques and can be readily propagated to high titer. Electron microscopy shows that Emalyn virions have a siphoviral morphotype with a long (300 nm) flexible tail, and a prolate head, 50 nm wide and 100 nm long.

The Emalyn genome was sequenced using an Illumina MiSeq with 140 bp single-end runs. Reads were assembled using Newbler to give a single major contig of 43,982 bp with 417-fold coverage. The viral genome has defined ends with 10 base 3' extensions (5'-CGGTAGGCTT), and an overall G+C content of 61.2%, somewhat lower than that of *Gordonia terrae* (67.8%). Sixty-six predicted protein coding genes were identified using Glimmer and Genemark (9, 10), 31 of which could be assigned putative functions using BLASTp, HHpred, and Phamerator (11, 12). No tRNA genes were identified.

Emalyn has extensive nucleotide sequence similarity (87% identity) to the previously described *Gordonia* phage GTE2 (3) spanning 95% of their genome length. The two phages share the organization of rightwards-transcribed virion structure and assembly genes in their left arms, followed by a lysis cassette, and nonstructural genes transcribed both leftwards and rightwards in the right genome parts.

Curiously, the lysis cassette includes two separate genes encoding endolysin activities found as domains within a single protein in the mycobacteriophages (e.g., Myrna gp243); Emalyn gp19 and

gp20 encode putative peptidase and glycoside hydrolase domains, respectively. Emalyn gp21 has two predicted transmembrane domains and is a holin candidate, supported by the finding that Emalyn gp21 homologues are found associated with endolysin genes in unrelated *Gordonia* phages. However, Emalyn gp22 has four predicted membrane spanning domains and may also be involved in lysis. Curiously, Emalyn encodes a putative lysin B protein (gp44), and homologues have been characterized previously only in the mycobacteriophages (13, 14). Furthermore, the lysin B gene (44) is atypically unlinked from the other lysis components, and sits in the right part of the genome downstream of a DNA polymerase I gene (42). The GTE2 homologue (gp37) is more distantly related to Emalyn gp44 (75% aa identity) than the DNA polymerase I proteins (95.6% aa identity). Other Emalyn-encoded proteins of note are RecA (gp57), two helicases (gp41, gp52), ThyX (gp46), and a kinase (gp50).

Accession number(s). The Emalyn genome is available from GenBank under accession number [KU963260](https://www.ncbi.nlm.nih.gov/nuclseq/KU963260).

FUNDING INFORMATION

This work, including the efforts of Graham F. Hatfull, was funded by Howard Hughes Medical Institute (HHMI) (54308198).

REFERENCES

- De los Reyes FL, III, Rothauszky D, Raskin L. 2002. Microbial community structures in foaming and nonfoaming full-scale wastewater treatment plants. *Water Environ Res* 74:437–449. <http://dx.doi.org/10.2175/106143002X140233>.
- Petrovski S, Seviour RJ, Tillett D. 2011. Prevention of *Gordonia* and *Nocardia* stabilized foam formation by using bacteriophage GTE7. *Appl Environ Microbiol* 77:7864–7867. <http://dx.doi.org/10.1128/AEM.05692-11>.
- Petrovski S, Seviour RJ, Tillett D. 2011. Characterization of the genome of the polyvalent lytic bacteriophage GTE2, which has potential for biocontrol of *Gordonia*-, *Rhodococcus*-, and *Nocardia*-stabilized foams in activated sludge plants. *Appl Environ Microbiol* 77:3923–3929. <http://dx.doi.org/10.1128/AEM.00025-11>.
- Petrovski S, Tillett D, Seviour RJ. 2012. Genome sequences and characterization of the related *Gordonia* phages GTE5 and GRU1 and their use as potential biocontrol agents. *Appl Environ Microbiol* 78:42–47. <http://dx.doi.org/10.1128/AEM.05584-11>.

5. 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. <http://dx.doi.org/10.1371/journal.pone.0134512>.
6. Liu M, Gill JJ, Young R, Summer EJ. 2015. Bacteriophages of wastewater foaming-associated filamentous *Gordonia* reduce host levels in raw activated sludge. *Sci Rep* 5:13754. <http://dx.doi.org/10.1038/srep13754>.
7. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs WR, Hendrix RW, Lawrence JG, Hatfull GF; Science Education Alliance Phage Hunters Advancing Enomics and Evolutionary Science, Phage Hunters Integrating Research and Education, Mycobacterial Genetics Course. 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife* 4:e06416.
8. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SC, 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:e01051-01013. <http://dx.doi.org/10.1128/mBio.01051-13>.
9. Borodovsky M, McIninch J. 1993. GeneMark: parallel gene recognition for both DNA strands. *Comput Chem* 17:123–133. [http://dx.doi.org/10.1016/0097-8485\(93\)85004-V](http://dx.doi.org/10.1016/0097-8485(93)85004-V).
10. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 27:4636–4641. <http://dx.doi.org/10.1093/nar/27.23.4636>.
11. 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. <http://dx.doi.org/10.1186/1471-2105-12-395>.
12. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <http://dx.doi.org/10.1093/nar/gki408>.
13. Payne K, Sun Q, Sacchettini J, Hatfull GF. 2009. Mycobacteriophage lysin B is a novel mycolylarabinogalactan esterase. *Mol Microbiol* 73:367–381. <http://dx.doi.org/10.1111/j.1365-2958.2009.06775.x>.
14. Gil F, Grzegorzewicz AE, Catalão MJ, Vital J, McNeil MR, Pimentel M. 2010. Mycobacteriophage Ms6 LysB specifically targets the outer membrane of *Mycobacterium smegmatis*. *Microbiology* 156:1497–1504. <http://dx.doi.org/10.1099/mic.0.032821-0>.