Genome Sequences of *Gordonia terrae* Phages Attis and SoilAssassin

Welkin H. Pope, Daniel N. Biery, Zachary T. Huff, Amy B. Huynh, William M. McFadden, Julia S. Mouat, Scott E. Schneiderman, Hannah Song, Leah E. Szpak, Melanie S. Umbaugh, Brian A. German, Jill E. McDonnell, Nadia Mezghani, Claire E. Schafer, Paige K. Thompson, Megan C. Ulbrich, Victor J. Yu, 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

Attis and SoilAssassin are two closely related bacteriophages isolated on *Gordonia terrae* 3612 from separate soil samples in Pittsburgh, PA. The Attis and SoilAssassin genomes are 47,881 bp and 47,880 bp, respectively, and have 74 predicted protein-coding genes, including toxin-antitoxin systems, but no tRNAs.

*Gordonia* spp. are implicated in foaming of sludge in wastewater treatment plants and are identified as opportunistic pathogens in hospital infections (1–4). Seventeen bacteriophages of *Gordonia* have been isolated, sequenced, and deposited in GenBank (5–9). It is unclear if the phages’ genomic relationships reflect those of other phages of the phylum Actinobacteria, notably those of *Mycobacterium smegmatis* mc²155 whose phages exhibit a continuum of genetic diversity (10–16). The Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program is a course-based research experience in which undergraduates are immersed in research, using phage isolation and bioinformatics as a method to fuse authentic research and education (17). SEA-PHAGES has recently expanded its range of hosts for phage isolation to *Gordonia terrae* 3612.

Attis and SoilAssassin were isolated from separate soil samples at the University of Pittsburgh through direct plating of filtered soil extracts on *G. terrae* 3612. The phages were plaque purified and electron microscopy revealed that both phages have long noncontractile tails and isometric heads. DNA was isolated, and sequenced using Illumina MiSeq technology with 140 bp single-end reads. Reams were assembled using Newbler into major contigs for each phage of 47,881 bp and 47,880 bp with 949-fold and 932-fold coverages for Attis and SoilAssassin, respectively. Both have G+C % content of 66.8%, and discrete genome ends with 11-base single-stranded 3′ extensions (5′-TACCAGGGGGA). BLASTn alignment of the two genomes shows that they differ by only a single base substitution and single bp insertion in Attis. Protein-coding genes were predicted using Glimmer (18), GeneMark (19), DNA Master (http://cobamide2.bio.pitt.edu), and PhamMaster (20), and functions were assigned to 35 of the 74 genes in each genome using BLAST (21) and HHPred (22, 23) alignment against the publicly available databases GenBank, the Protein DataBase, and pFamA. Predicted functions include those for virion structure genes, a tyrosine integrase and immunity repressor, a RecET recombination system, RusA resolvase, and two HNH endonucleases.

Attis and SoilAssassin form plaques with distinct morphologies. SoilAssassin plaques are large (4-mm diameter), whereas Attis forms predominantly smaller plaques (1-mm diameter) although a few larger plaques are also observed. We predict that the single bp deletion at coordinate 25,254 of SoilAssassin in a putative minor tail protein gene is responsible for the plaque size difference. We note that Attis 29 likely corresponds to the parental form of the gene as the entire coding sequence is related to genes in several other distant phages. This relationship is reminiscent of the side tail fiber gene frameshift mutation in PaPa strains of phage lambda that facilitate large plaque formation relative to its Ur-lambda parent, which adsorbs to cells more rapidly than lambda PaPa thus reducing plaque size (24). Attis and SoilAssassin 29 genes are located at the 3′ end of the virion structure and assembly operons, consistent with encoding tail fiber proteins.

Both phages encode putative tyrosine integrase and repressor genes near the center of their genomes, and are predicted to integrate into a tRNA⁶⁰⁰ gene corresponding to GTR9_RS02590 in *Gordonia* sp. KTR9.

**Nucleotide sequence accession numbers.** The Attis and SoilAssassin genomes are available from GenBank under accession numbers KU963247 and KU963246, respectively.

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

REFERENCES

1. Blaschke AJ, Bender J, Byington CL, Korgenski K, Daly J, Petti CA, Pavia AT, Ampofo K. 2007. *Gordonia* species: emerging pathogens in pediatric patients that are identified by 16S ribosomal RNA gene sequencing. Clin Infect Dis 45:483–486. http://dx.doi.org/10.1086/520018.

2. De los Reyes FL, III, Raskin L. 2002. Role of filamentous microorganisms in activated sludge foaming: relationship of mycolata levels to foaming initiation and stability. Water Res 36:445–459. http://dx.doi.org/10.1016/S0043-1354(01)00227-5.
3. De los Reyes FL, III, Rothauszky D, Raskin L. 2002. Microbial community structures in foaming and nonfoaming full-scale wastewater treatment plants. Water Environment Res 74:437–449. http://dx.doi.org/10.2175/106143002X140233.

4. Grisold AJ, Roll P, Hoennig M, Feierl G, Vicenzi-Moser R, Marth E. 2007. Isolation of Gordonia terrae from a patient with catheter-related bacteremia. J Med Microbiol 56:1687–1688. http://dx.doi.org/10.1099/jmm.0.47388-0.

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 biofilm control. 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-associating filamentous Gordonia reduce host levels in raw activated sludge. Sci Rep 5:31754. http://dx.doi.org/10.1038/srep31754.

7. Petrovski S, Seviour RJ, Tillett D. 2011. Prevention of Gordonia and Nocardioides stabilized foam formation by using bacteriophage GETE. Appl Environ Microbiol 77:7864–7867. http://dx.doi.org/10.1128/AEM.05692-11.

8. 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 Nocardioides-stabilized foams in activated sludge plants. Appl Environ Microbiol 77:3923–3929. http://dx.doi.org/10.1128/AEM.00023-11.

9. Petrovski S, Tillett D, Seviour RJ. 2012. Genome sequences and characterization of the related Gordonia phages GET5 and GRU1 and their use as potential biocontrol agents. Appl Environ Microbiol 78:42–47. http://dx.doi.org/10.1128/AEM.05584-11.

10. Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, Weber RJ, Patel MG, Germane KL, Edgar RH, Hoyte NN, Bowman CA, Tantoco AT, Paladín EC, Myers MS, Smith AI, Grace MS, Pham TT, O’Brien MB, Vogelsberger AM, Hryckowian AJ, Wynalek JL, Doniskeller 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. http://dx.doi.org/10.1016/j.jmb.2010.01.011.

11. Jacobs-Sera D, Marinelli LJ, Bowman C, Broussard GW, Guerrero Bustamante C, Boeing MM, Petrova ZO, Dedrick RM, Pope WH, Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science Sea-Phages Program, Modlin RL, Hendrix RW, Hatfull GF. 2012. On the nature of mycobacteriophage diversity and host preference. Virology 434:187–201.

12. Pope WH, Anders KR, Baird M, Bowman CA, Boeing MM, Broussard GW, Chow T, Class KL, Cooper S, Corney EA, Dejong RJ, Delesalle VA, Deng L, Dunbar D, Edgington NF, Ferreira CM, Weston Hafer JK, Hartzog GA, Hatherill JR, Hughes LE, Iapko K, Krukonis GP, Meier CG, Monti DL, Olm MR, Page ST, Peebles CL, Rinehart CA, Rubin MR, Russell DA, Sanders ER, Schoer M, Shaffer CD, Wherley J, Vazquez E, Yuan H, Zhang D, Cresawn SG, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2014. Cluster M mycobacteriophages Bongo, PegLeg, and Rey with unusually large repertoires of RNA isotypes. J Virol 88:2461–2480. http://dx.doi.org/10.1128/JVI.01336-13.

13. 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 Genomics and Evolutionary Science, Phage Hunters Integrating Research and Education, >Myco- bacteriologic Genetics Course. 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife 4:e0416. http://dx.doi.org/10.7554/eLife.0416.

14. Pope WH, Ferreira CM, Jacobs-Sera D, Benjamin RC, Davis AJ, Dejong RJ, Elgin SC, Guilfoile FR, Forshy MH, Harris AD, Harvey SE, Hughes LE, Hynes PM, Jackson AS, Jalal MD, MacMurray EA, Manley CM, McDonough MJ, Mosier JL, Osterbert LM, Rabinowitz HS, Rhyasen CN, Russell DA, Saha MS, Shaffer CD, Simon SE, Sims EF, Tovar IG, Weisser EG, Wertz JT, Weston-Hafer KA, Williamson KE, Zhang R, Cresawn SG, Jain P, Piuri M, Jacobs WR, Jr, Hendrix RW, Hatfull GF. 2011. Cluster K mycobacteriophages: insights into the evolutionary origins of mycobacteriophage TM4. PLoS One 6:e26750. http://dx.doi.org/10.1371/journal.pone.0026750.

15. Pope WH, Jacobs-Sera D, Best AA, Broussard GW, Connerly PL, Dedrick RM, Kremer TA, Offner S, Ogiefo AH, Pizzorno MC, Rockenbach K, Russell DA, Stowe EL, Stukay J, Thibault SA, Conway JF, Hendrix RW, Hatfull GF. 2013. Cluster J mycobacteriophages: introns in capsid and tail genes. PLoS One 8:e69273. http://dx.doi.org/10.1371/journal.pone.0069273.

16. Pope WH, Jacobs-Sera D, Russell DA, Peebles CL, Al-Atraque Z, Alcoser TA, Alexander LM, Alfano MB, Alford ST, Amy NE, Anderson MD, Anderson AG, Ang AA, Ares M, Jr, Barber AJ, Barker LP, Barrett JM, Barshop WD, Bauere BM, Bayles DM, Belfield KL, Best AA, Borjon A, Jr, Bowman CA, Boyer CA, Bradley KW, Bradley VA, Browning LN, Budwal K, Busby KN, Campbell JW, Campbell AM, Carey A, Caruso SM, Chee RD, Cockburn CL, Cohen LB, Corajon JD, Cresawn SG, Davis KR, Deng L, Denver DR, Dixon BR, Ekram S, Elgin SC, Engels AE, English BE, Erb ML, Estrada C, Filliger LZ. 2011. Expanding the diversity of mycobacteriophages: insights into genome architecture and evolution. PLoS One 6:e16329. http://dx.doi.org/10.1371/journal.pone.0016329.

17. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, Dejong BJ, Denver DR, Dunbar D, Elgin SF, Findlay AM, Gissendanner CR, Golebiewska UP, Guild N, Hartsington J, Holwell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenweig F, Rubin MR, Saha MS, Sadowsz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradford 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 undergraduates. MBio 5:e00159-01013. http://dx.doi.org/10.1128/mBio.00159-13.

18. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinformatics 23:673–679. http://dx.doi.org/10.1093/bioinformatics/btm009.

19. Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotic, eukaryotic and viruses. Nucleic Acids Res 33:W451–W454. http://dx.doi.org/10.1093/nar/gki467.

20. 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.

21. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80060-2.

22. Riemert M, Bieger A, Hauser A, Söding J. 2012. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9:173–175. http://dx.doi.org/10.1038/nmeth.1818.

23. Söding J, Bieger 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.

24. Hendrix RW, Duda RL. 1992. Bacteriophage Lambda PaPa: not the mother of all lambda phages. Science 258:1145–1148. http://dx.doi.org/10.1126/science.1439823.