# Genome Sequences of *Gordonia terrae* Phages Benczkowski14 and Katyusha

Welkin H. Pope, Matthew S. Benczkowski, Daryn E. Green, Melina Hwang, Bryan Kennedy, Bradley Kocak, Ellen Kruczek, Leon Lin, Matthew L. Moretti, Faith L. Onelangsy, Nadia Mezghani, Katherine A. Milliken, Chelsea L. Toner, Paige K. Thompson, Megan C. Ulbrich, 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

**Bacteriophages Katyusha and Benczkowski14 are newly isolated phages that infect *Gordonia terrae* 3612. Both have siphoviral morphologies with isometric heads and long tails (500 nm). The genomes are 75,380 bp long and closely related, and the tape measure genes (9 kbp) are among the largest to be identified.**

Address correspondence to Welkin H. Pope, welkin@pitt.edu.

*G*ordonia spp. are common soil bacteria and are also associated with wastewater treatment plants (1). Several phages of *Gordonia* hosts have been isolated and sequenced, all of which have siphoviral morphologies, and many of which have unusually long tails (2–6). The Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program offers a course-based undergraduate research experience for advancing our understanding of bacteriophage diversity (7, 8).

Phages Benczkowski14 and Katyusha were isolated by enrichment of soil samples collected from Pittsburgh, PA, using *Gordonia terrae* 3612 as a host. Following plaque purification, amplification, and DNA extraction, the genomes were sequenced using Illumina MiSeq. Single-end 140-bp reads were assembled using Newbler, and both genomes assembled into a single major contig, with average coverages of 944-fold and 1,419-fold for Benczkowski14 and Katyusha, respectively. The genome coverage indicates that both have 1,172-bp direct terminal repeats and genome lengths of 75,380 bp. The genomes are almost identical and differ by just three single-nucleotide substitutions, but they are not closely related to previously reported phage genomes.

Protein-coding genes were predicted using Glimmer (9) and GeneMark (10), and functional assignments were made using BLASTP (11), HHpred (12), and Phamerator (13). Each genome has 99 putative protein-coding genes but no tRNA genes, and functional assignments could be made to fewer than 30% of the predicted genes. These include the virion structure and assembly genes, helicases, Holliday junction resolvases, HNH endonucleases, and WhiB regulators. Both genomes also encode a CobT-like protein.

Katyusha and Benczkowski14 have a strikingly large tape measure protein gene (9,084 bp), consistent with the length of the virion tails (approximately 500 nm). These are among the longest phage genes identified, and only slightly shorter than the tape measure protein gene of *Gordonia* phage GMA7 (9,141 bp [2]). Although GMA7 does not share extensive nucleotide sequence similarity with Katyusha and Benczkowski14, many of the virion structural proteins are related at the amino acid sequence level, including the portal, capsid maturation protease, major capsid subunit, major tail subunit, the tape measure protein, and some of the minor tail genes. The tape measure proteins share 50% amino acid identity with the tape measure protein of GMA7 and contain putative lytic transglycosylase domains, as reported for many mycobacteriophage tape measure proteins (14).

Benczkowski14, Katyusha, and GMA7 each have two genes located downstream of the virion structural gene operon coding for lysis functions, with one coding for a muramidase and a second coding for a peptidase. All three genomes also have two or more closely linked genes encoding putative protein products, at least one of which is anticipated to provide the holin function. We did not identify either integrase or putative repressor genes, which is consistent with a strictly lytic lifestyle.

**Nucleotide sequence accession numbers.** The genomes of Benczkowski14 and Katyusha are available from GenBank under the accession numbers KU963262 and KU963258, respectively.

## REFERENCES

1. **De los Reyes III FL, 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.
2. **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.
3. **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.
4. **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.

5. **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.

6. **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.

7. **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**:e01051-13. http://dx.doi.org/10.1128/mBio.01051-13.

8. **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, Mycobacterial Genetics Course.** 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife **4**:e06416.

9. **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.

10. **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.

11. **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)80360-2.

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. **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.

14. **Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, Jacobs-Sera D, Falbo J, Gross J, Pannunzio NR, Brucker W, Kumar V, Kandasamy J, Keenan L, Bardarov S, Kriakov J, Lawrence JG, Jacobs WR, Hendrix RW, Hatfull GF.** 2003. Origins of highly mosaic mycobacteriophage genomes. Cell **113**:171–182. http://dx.doi.org/10.1016/S0092-8674(03)00233-2.