1-1-2018

Genome Sequences of Four Cluster P Mycobacteriophages

Erin L. Doyle
Doane University

Christy L. Fillman
University of Colorado Boulder

Nathan S. Reyna
Ouachita Baptist University, reynan@obu.edu

Ruth C. Plymale
Ouachita Baptist University, plymaler@obu.edu

Deborah M. Tobiason
Carthage College

See next page for additional authors

Recommended Citation

Doyle, Erin L., Fillman, Christy L., Reyna, Nathan S., Plymale, Ruth, et al. "Genome Sequences of Four Cluster P Mycobacteriophages," Genome Announcements, American Society for Microbiology, 6:2 (2018) Jan 11. doi: 10.1128/genomeA.01101-17.

Authors

Erin L. Doyle, Christy L. Fillman, Nathan S. Reyna, Ruth C. Plymale, Deborah M. Tobiason, Daniel E. Westholm, Jonathan L. Askins, Brittany P. Backus, Ashlynn C. Baker, Harrison S. Ballard, Paul J. Bisesi, Logan Bond, Deanna Byrnes, Hannah Carlstedt, Kinnon S. Dodson, Megan J. Fallert, Kyla J. Foster, Daniel N. Games, Tristan R. Grams, Nancy A. Guild, Autumn Hurd, Nicholas Iwata, Cassidy R. Kepler, Lucinda R. Krenzke, Kelly Luekens, Jackie Lewis, Cali McEntee, Justin C. McGee, Noah Nalley, Jade Prochaska, Reid G. Rogers, Jessica B. Schipper, and Kelsey Snyder

AMERICAN SOCIETY FOR MICROBIOLOGY genomeAnnouncements™

VIRUSES

Genome Sequences of Four Cluster P Mycobacteriophages

Erin L. Doyle,[a] Christy L. Fillman,[b] Nathan S. Reyna,[c] Deborah M. Tobiason,[d] Daniel E. Westholm,[e] Jonathan L. Askins,[c] Brittany P. Backus,[c] Ashlynn C. Baker,[a] Harrison S. Ballard,[c] Paul J. Bisesi,[b] Logan Bond,[c] Deanna Byrnes,[d] Hannah Carlstedt,[a] Kinnon S. Dodson,[c] Megan J. Fallert,[d] Kyla J. Foster,[b] Daniel N. Games,[c] Tristan R. Grams,[d] Nancy A. Guild,[b] Autumn Hurd,[a] Nicholas Iwata,[a] Cassidy R. Kepler,[a] Lucinda R. Krenzke,[d] Kelly Luekens,[a] Jackie Lewis,[a] Cali McEntee,[a] Justin C. McGee,[c] Noah Nalley,[c] Ruth C. Plymale,[c] Jade Prochaska,[a] Reid G. Rogers,[c] Jessica B. Schipper,[b] Kelsey Snyder,[a] Kali Uhrich,[a] Chelsey D. Vermillion,[c] Sarah K. Vickers,[c] Meredyth D. Wenta,[d] Tyler Z. Yates,[b] Chas F. Young,[d] Ty H. Stoner,[f] Welkin H. Pope,[f] Deborah Jacobs-Sera,[f] Rebecca A. Garlena,[f] Daniel A. Russell,[f] Steven G. Cresawn,[g] Graham F. Hatfull[f]

[a]Biology Department, Doane University, Crete, Nebraska, USA
[b]Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado, USA
[c]Biology Department, Ouachita Baptist University, Arkadelphia, Arkansas, USA
[d]Biology Department, Carthage College, Kenosha, Wisconsin, USA
[e]Biology Department, The College of St. Scholastica, Duluth, Minnesota, USA
[f]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[g]Department of Biology, James Madison University, Harrisonburg, Virginia, USA

ABSTRACT Four bacteriophages infecting *Mycobacterium smegmatis* mc^{2}155 (three belonging to subcluster P1 and one belonging to subcluster P2) were isolated from soil and sequenced. All four phages are similar in the left arm of their genomes, but the P2 phage differs in the right arm. All four genomes contain features of temperate phages.

A large collection of over 1,400 sequenced bacteriophages infecting *Mycobacterium smegmatis* mc^{2}155 reveals that they have considerable genetic diversity (1). Phages grouped in clusters N and P are temperate but use an unusual integration-dependent immunity system, in which *attP* is positioned within the repressor gene (2); genome integration results in the truncation of the repressor, removal of a C-terminal degradation tag, and expression of stable repressor protein.

Mycobacteriophages Bartholomew, Bogie, Ksquared, and Tortellini were isolated from soil samples from Kenosha, Wisconsin, USA; Camden, Arkansas, USA; Crete, Nebraska, USA; and Boulder, Colorado, USA, respectively, via enrichment cultures and plaque purification. Bartholomew and Bogie produced plaques with large halos and clear centers, while Ksquared and Tortellini produced turbid plaques. All four phages have siphoviral morphologies.

Purified DNA was sequenced using the Illumina MiSeq platform, and 140-bp single-end reads were assembled into a major contig for each phage using Newbler, with at least 150-fold coverage. Genome lengths range from 46,484 bp to 49,658 bp (Table 1), and G+C% contents range from 65.8% to 67.2%, similar to the host bacterium (67.4%). All four genomes have defined ends with 12 base 3′ single-strand DNA extensions (Table 1). Bartholomew, Bogie, and Ksquared are closely related to each other and to other subcluster P1 phages at the nucleotide sequence level. Tortellini is closely related ($>$80% nucleotide identity) to these phages at the leftmost 40% of the genome but diverges in the rightmost 60% of the genome ($<$22% nucleotide identity). Tortellini is thus assigned as the first member of subcluster P2. Interestingly, although the rightmost 26-kbp segment of Tortellini has diverged considerably from subcluster P1 phages, the P1 phages have extensive sequence similarity to cluster I phages in this region (3). Genomes were auto-annotated using Glimmer (4) and GeneMark (5) and

Received 7 September 2017 **Accepted** 6 November 2017 **Published** 11 January 2018

Citation Doyle EL, Fillman CL, Reyna NS, Tobiason DM, Westholm DE, Askins JL, Backus BP, Baker AC, Ballard HS, Bisesi PJ, Bond L, Byrnes D, Carlstedt H, Dodson KS, Fallert MJ, Foster KJ, Games DN, Grams TR, Guild NA, Hurd A, Iwata N, Kepler CR, Krenzke LR, Luekens K, Lewis J, McEntee C, McGee JC, Nalley N, Plymale RC, Prochaska J, Rogers RG, Schipper JB, Snyder K, Uhrich K, Vermillion CD, Vickers SK, Wenta MD, Yates TZ, Young CF, Stoner TH, Pope WH, Jacobs-Sera D, Garlena RA, Russell DA, Cresawn SG, Hatfull GF. 2018. Genome sequences of four cluster P mycobacteriophages. Genome Announc 6:e01101-17. https://doi.org/10.1128/genomeA.01101-17.

Address correspondence to Graham F. Hatfull, gfh@pitt.edu.

TABLE 1 Genomic information for cluster P bacteriophages

Phage name	Accession no.	Genome size (bp)	GC%	No. of genes	ssDNA[a] termini
Bartholomew	MF140399	46,484	67.2	77	CCCGCCCCCCGA
Bogie	MF133446	48,639	66.9	81	CCCGCCCCCCGA
Ksquared	MF281061	48,699	67.1	80	CCTGCCGCCCGA
Tortellini	KX648391	49,658	65.8	76	CCTGCCGCCCGC

[a] ssDNA, single-stranded DNA.

refined by manual inspection; the number of protein-coding genes ranges from 76 to 81 (Table 1). Functions were assigned using Phamerator (6), BLASTp (7), and HHpred (8). No tRNAs are present in any of the genomes.

All four phages have features of temperate phages, including a tyrosine integrase, an immunity repressor, and an antirepressor. The three P1 phages contain integration-dependent immunity systems based upon the location of the *attP* site within their immunity repressor genes, overlapping a host tRNA-Thr$^{(CGT)}$ gene, and SsrA-like degradation tags at the ends of the integrases and repressors (2). Similar systems are common in cluster P phages (3, 9). Tortellini is unusual among cluster P phages in that it does not have the characteristics of the integration-dependent immunity systems, and we could not identify a plausible *attP* site. A putative programmed frameshift is predicted to be involved in the expression of the tail assembly chaperone genes located upstream of the tape measure gene in all four phages. The three P1 phages have lysis cassettes with putative lysin A, lysin B, and holin genes, but Tortellini lacks a lysin B gene. In addition, the P1 phages code for a RecE/RecT system, an FtsK DNA translocase, and a RusA-like Holliday junction resolvase, which are not encoded in Tortellini.

Accession number(s). GenBank accession numbers are shown in Table 1.

ACKNOWLEDGMENTS

We thank the SEA-PHAGES program for support.

This work was supported by Howard Hughes Medical Institute grant 54308198 to G.F.H. Undergraduate research at Ouachita Baptist University is supported by grants from the National Center for Research Resources (P20RR016460) through the National Institute of General Medical Sciences and the National Institutes of Health (P20GM103429), and by National Science Foundation EPSCoR grant IIA-1457888.

REFERENCES

1. Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, 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.
2. Broussard GW, Oldfield LM, Villanueva VM, Lunt BL, Shine EE, Hatfull GF. 2013. Integration-dependent bacteriophage immunity provides insights into the evolution of genetic switches. Mol Cell 49:237–248. https://doi.org/10.1016/j.molcel.2012.11.012.
3. 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. https://doi.org/10.7554/eLife.06416.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. Hatfull GF, Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) Program, KwaZulu-Natal Research Institute for Tuberculosis and HIV (K-RITH) Mycobacterial Genetics Course, University of California–Los Angeles Research Immersion Laboratory in Virology, Phage Hunters Integrating Research and Education (PHIRE) Program. 2016. Complete genome sequences of 61 mycobacteriophages. Genome Announc 4(4):e00389-16. https://doi.org/10.1128/genomeA.00389-16.