11-28-2023

Genome Sequence and Annotation of the B3 Mycobacteriophage Phayeta

Emily Bishop
Coastal Carolina University

Warren Earley
Coastal Carolina University

Alexandra Greco
Coastal Carolina University

Emma Hofseth
Coastal Carolina University

Emma Kinerson
Coastal Carolina University

See next page for additional authors

Recommended Citation

Bishop, E., Earley, W., Greco, A., Hofseth, E., Kinerson, E., Lafayette, B., Llanot-Arocho, N., Mazen, B., Cevasco, M., & Williams, D.C. (2023). Genome sequence and annotation of the B3 mycobacteriophage Phayeta. Microbiology Resource Announcements, e00915-23. https://doi.org/10.1128/MRA.00915-23. Available at https://digitalcommons.coastal.edu/biology/6/.

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Bacteriophages | Announcement

Genome sequence and annotation of the B3 mycobacteriophage Phayeta

Emily Bishop,[1] Warren Earley,[1] Alexandra Greco,[1] Emma Hofseth,[1] Emma Kinerson,[1] Brandon Lafayette,[1] Nestor Llanot-Arocho,[1] Brittney Mazen,[1] Megan Cevasco,[1] Daniel C. Williams[1]

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT Mycobacteriophage Phayeta was extracted from soil near Myrtle Beach, South Carolina using *Mycobacterium smegmatis* as a host. Annotation of the 68,700 base-pair circularly permuted genome identified 104 predicted protein-encoding genes, 34 of which have functional assignments.

KEYWORDS Bacteriophage

Mycobacterium smegmatis is a genetically tractable model of *Mycobacterium tuberculosis* and phages that infect *M. smegmatis* have the potential to be therapeutically useful (1, 2). Mycobacteriophage Phayeta was directly isolated from a soil sample taken near Myrtle Beach, South Carolina (Global Positioning Coordinates 33.75493N, 78.90047 W), using standard protocols (3). The soil sample was washed in 7H9 liquid media, filtered (0.22 um), and 500 uL of the filtrate mixed with 250 mL of saturated *M. smegmatis* mc^2155 for 15 minutes before being plated in top agar and incubated at 37°C for 48 hours. Phayeta produced small clear plaques that are characteristic of non-temperate phages (Fig. 1A). Plaque size was measured with ImageJ and ranged from 0.5 mm to 1.7 mm with an average of 1.0 ± 0.2 mm (mean ± SD, n = 50) (4, 5). Negative-staining transmission electron microscopy of Phayeta (Fig. 1B) revealed a siphovirus morphology consisting of an isometric capsid (diameter = 76.6 ± 3.6 nm) and a long flexible tail (274.7 ± 11.6, n = 9 viral particles).

Phayeta DNA was extracted from high-titer lysates using the Wizard DNA Clean-Up Kit (Promega) and analyzed by agarose gel electrophoresis. Intact DNA was prepared with the Ultra II Library Kit (New England Biolabs) and sequenced on the Illumina MiSeq platform (v3 reagents). A total of 446,915 single-end 150 base reads provided 278× coverage of the Phayeta genome. Sequence reads were assembled and verified using Newbler (v2.9) and Consed (v29) (6). The genome is 68,700 base-pairs long, has 67.5% GC content, and contains circularly permuted genomic ends. Comparison of gene nucleotide similarity of phages within the Actinobacteriophage Database (phagesdb.org), placed Phayeta in subcluster B3 (7). Among this cluster, Phayeta shares the greatest gene content similarity with Casbah (94.6%) and Kronus (91.7%).

Initial draft annotation of Phayeta was done using the GenMark (v2.5), Glimmer (v3.02), and ARAGORN algorithms within DNA Master (v5.23.56) using the bacterial, archaeal, and plant plastid code (8–10). Subsequent manual annotations employed Phamerator (v7), Starterator (v1.3), and PECAAN (11). Both PhagesDB and NCBI nonredundant databases were searched for sequence-based homology comparison via BLAST (7, 12). Predicted protein structure comparisons used HHPRED (v3.18) to search the PDB mmCIF70, Pfam-A, and NCBI Conserved Domain databases using default parameters. A total of 104 protein-encoding genes were identified, 34 of which have a functional assignment.

Editor John J. Dennehy, Queens College, Queens, New York, USA

Address correspondence to Daniel C. Williams, dwilliams@coastal.edu.

The authors declare no conflict of interest.

Received 28 September 2023
Accepted 29 September 2023

Published 28 November 2023

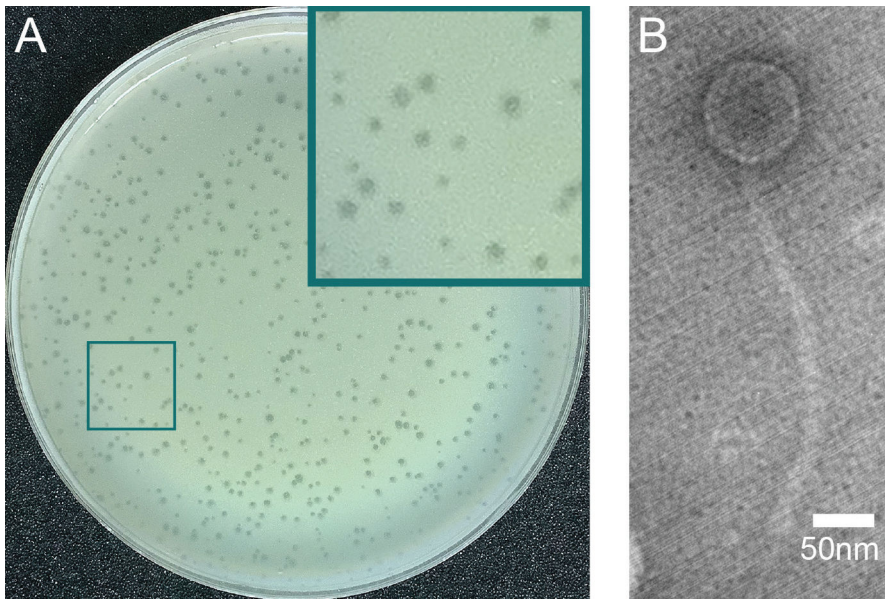

FIG 1 The mycobacteriophage Phayeta. (A) Plaque assay. Purified samples of Phayeta (10 ul) were mixed with 250 ul of saturated *M. smegmatis* MC^2155, then combined with 3 mL of molten 7H9 top-agar and poured on 90 mm plates. Incubation at 37°C for 48 hours resulted in the production of clear small plaques. (B) Transmission electron microscopy of Phayeta negatively stained with 1% uranyl acetate and imaged using a JEOL JEM 1230 at 100 kV.

Most genes on the left arm of Phayeta encode structural or assembly proteins with the right arm containing genes involved in DNA replication. Consistent with a non-temperate lifestyle, Phayeta lacks genes encoding an identifiable integrase or repressor. Both Lysin A and Lysin B are present. Although homology searching failed to identify a holin-encoding gene, a small cluster of transmembrane-segment-containing genes (gp17-gp20) have sequence and syntenic similarities with genes of *Gordonia* phages that are a predicted holin cassette (13). There are five minor tail proteins, including gp33 which shows full-length similarity to Corofin gp33. Both these phages have a single nucleotide indel polymorphism in this gene but at different locations suggesting distinct mutational events.

ACKNOWLEDGMENTS

We thank Dan Russell and Ching-Chung Ko of the Pittsburch Bacteriophage Institute for sequencing and genome assembly, Julian Smith III and Victoria Frost of Winthrop University for their electron microscopy service, and the Howard Hughes Medical Institute SEA-PHAGES program and Coastal Carolina University Biology Department for support.

AUTHOR AFFILIATION

[1]Department of Biology, Coastal Carolina University, Conway, South Carolina, USA

AUTHOR ORCIDs

Megan Cevasco http://orcid.org/0000-0003-2027-2225
Daniel C. Williams http://orcid.org/0000-0002-7900-480X

DATA AVAILABILITY

Phayeta is available at GenBank with Accession No. OR159655 and Sequence Read Archive No. SRX20630268.

REFERENCES

1. Shiloh MU, Champion PAD. 2010. To catch a killer. what can mycobacterial models teach us about *Mycobacterium tuberculosis* pathogenesis? Curr Opin Microbiol 13:86–92. https://doi.org/10.1016/j.mib.2009.11.006
2. Hatfull GF. 2022. Mycobacteriophages: from petri dish to patient. PLoS Pathog 18:e1010602. https://doi.org/10.1371/journal.ppat.1010602
3. Poxleitner M, Pope WH, Jacobs-Sera D, Sivanathan V, Hatfull GF. 2018. SEA-PHAGES phage discovery guide. Available from: https://seaphagesphagediscoveryguide.helpdocsonline.com/home
4. Collins TJ. 2007. ImageJ for microscopy. Biotechniques 43:25–30. https://doi.org/10.2144/000112517
5. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods 9:676–682. https://doi.org/10.1038/nmeth.2019
6. Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes, p 109–125. In Clokie MRJ, AM Kropinski, R Lavigne (ed), Bacteriophages. Springer, New York, New York, NY. https://doi.org/10.1007/978-1-4939-7343-9
7. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711
8. 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
9. 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
10. 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
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. https://doi.org/10.1186/1471-2105-12-395
12. 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
13. Pollenz RS, Bland J, Pope WH. 2022. Bioinformatic characterization of endolysins and holin-like membrane proteins in the lysis cassette of phages that infect *Gordonia rubripertincta*. PLoS One 17:e0276603. https://doi.org/10.1371/journal.pone.0276603