Genome Sequences of Two *Gordonia rubripertincta* Cluster DJ Bacteriophages, Pherobrine and Burley

Aaron J. Adams,a Michael Bell,a Marie J. Bennett,a Umar N. Chaudhry,a Jennifer N. Daniels,a James E. Herman,a Sophia A. Oh,a Fedora Parra,a Cynthia R. Reagan,a Borna Zareiesfandabadi,a Marie P. Fogartya

aScience Department, Durham Technical Community College, Durham, North Carolina, USA

**ABSTRACT** Pherobrine and Burley are siphoviruses infecting *Gordonia rubripertincta*. Pherobrine has a 60,305-bp genome with 89 predicted protein-coding genes, and Burley has a 60,111-bp genome with 90 predicted protein-coding genes. Both phages are assigned to cluster DJ, where they share 78% gene content similarity with each other.

The isolation and genomic characterization of novel bacteriophages can support the development of therapeutics to address the global increase in antibiotic-resistant bacterial infections (1). Here, we report the isolation and genome sequence of two lytic bacteriophages, Pherobrine and Burley.

Pherobrine and Burley were extracted from soil samples collected in August 2021 in Durham (36.118889°N, 78.864722°W) and Cedar Grove (36.2209°N, 79.16892°W), North Carolina, respectively, using *Gordonia rubripertincta* NRRL B-16540 and standard isolation procedures (2). Briefly, the soil samples were washed with peptone-yeast extract-calcium (PYCa) liquid medium, the wash was filtered through 0.22-μm filters, and the filtrate was inoculated with *G. rubripertincta* and incubated with shaking at 30°C for 3 to 5 days. An aliquot of the culture was filtered and diluted, and dilutions were plated in top agar containing *G. rubripertincta* and incubated at 30°C. Both phages, which produced clear plaques with an ~1-mm-wide diameter after 3 days at 30°C, were purified through three rounds of plating. Lysates were negatively stained using 2% phosphotungstic acid (PTA) and imaged by transmission electron microscopy (TEM) to reveal siphovirus morphologies. The capsid diameter was 72 to 80 nm (n = 5 for each phage), while the average tail length was 240 to 264 nm (n = 5 for each phage) (Fig. 1).

DNA was extracted from Pherobrine and Burley using phenol-chloroform-isoamyl alcohol extraction (Sigma-Aldrich; catalog no. P2069), prepared for sequencing using the NEBNext Ultra II library kit, and sequenced using the Illumina MiSeq (v3 reagents) platform to yield 150-bp single-end reads that provided 1,543-fold genome coverage of Pherobrine and Burley.
360-fold genome coverage of Burley. The reads were assembled using Newbler v2.9 and checked for accuracy, coverage, and genomic termini using Consed v29 (3–5). The sequencing results and genome characteristics are listed in Table 1. Based on gene content similarity (GCS) of at least 35% to phages in the Actinobacteriophage database, phagesDB (6), both Pherobrine and Burley, which share 78% GCS, were assigned to phage cluster DJ.

The genomes of Pherobrine and Burley were auto-annotated using Glimmer v3.02 (7) and GeneMark v2.5p (8). Manual inspection and annotation refinement were carried out using DNA Master v5.23.6 (9), PECAAN (https://discover.kbrinsgd.org/), Phamerator (10), Starterator v1.2 (9), NCBI BLASTp (11), and HHpred v3.2 (PDB, CDD, pFamA, UniProt) (12). TMHMM v2.0 (13) and SOSUI v1.11 (14) were used for transmembrane domain prediction, and ARAGORN v1.2.38 (15) and tRNAscan-SE v2.0 (16) were used for tRNA prediction. Default parameters were used for all software. Pherobrine and Burley contain 89 and 90 genes, respectively, all of which are transcribed rightward. Genes encoding structural, assembly, and lysis functions are encoded on the left half of the genome, whereas DNA metabolism functions are encoded on the right arm of the genome. At least 12 membrane proteins were predicted in each phage, while no tRNA, immunity repressor, or integrase genes could be identified. Notably, and characteristic of cluster DJ phages, a region spanning ~7 kbp downstream of tape measure gene contains multiple sequence repeats that occur intergenically and upstream of ATG gene start sites, suggesting an involvement in gene regulation.

Data availability. Pherobrine is available at GenBank with accession no. ON970572 and Sequence Read Archive (SRA) with accession no. SRX14483240, and Burley is available at GenBank with accession no. ONS26971 and SRA with accession no. SRX15121737.

ACKNOWLEDGMENTS

We thank the HHMI SEA-PHAGES program for support and Daniel Russell and Rebecca Garlena for sequencing and assembling the genomes. The phages were isolated by undergraduates in fall 2021. We thank Catherine Ward for discovery course coinstruction and thank Kristen White at UNC-CH microscopy services laboratory for obtaining TEM micrographs. Richard Pollenz provided helpful feedback on the presence of intergenic repeats upstream of ATG gene start sites.

REFERENCES

1. Hatfull GF. 2022. Mycobacteriophages: from petri dish to patient. PLoS Pathog 18:e1010602. https://doi.org/10.1371/journal.ppat.1010602.
2. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull GF. 2018. HHMI SEA-PHAGES phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD. https://seaphages.org/blog/2018/09/05/sea-phages-phage-discovery-guide-online/.
3. Genivaldo GZ, Silva BE, Duttilh D, Matthews K, Elkins R, Schmieder EA, Dinsdale RAE. 2013. Combining de novo and reference-guided assembly with scaffold_builder. Source Code Biol Med 8:23.
4. Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. Bioinformatics 29:2936–2937. https://doi.org/10.1093/bioinformatics/bt315.
5. Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes. Methods Mol Biol 1681:109–125. https://doi.org/10.1007/978-1-4939-7343-9_9.
6. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711.
7. Delcher AL, Batke 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.
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. Pope WH, Jacobs-Sera D. 2018. Annotation of bacteriophage genome sequences using DNA Master: an overview. Methods Mol Bio 1681:217–229. https://doi.org/10.1007/978-1-4939-7343-9_16.
10. 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.
11. 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/10.1016/0022-2836(90)90320-2.
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. https://doi.org/10.1093/nar/gki408.
13. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a Markov model: application to complete genomes. J Mol Biol 305:567–580. https://doi.org/10.1006/jmbi.2000.4315.

14. Hirokawa T, Boon-Chieng S, Mitaku S. 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14:378–379. https://doi.org/10.1093/bioinformatics/14.4.378.

15. 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.

16. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.3.955.