



Genome Sequences of Three Cluster C Mycobacteriophages, Bipolarisk, Bread, and FudgeTart

Dane M. Bowder,^a Brandon W. Gannon,^a Kathryn J. Grint,^a Jason T. Iltz,^a Teryn M. Koch,^a Kaitlyn S. Mahnke,^a Brea D. Murnan,^a Alexandria M. Osborn,^a Danielle M. Schreiber,^a Grace L. Su,^a Jaime G. Troester,^a Erin L. Doyle^a

^aDepartment of Biology, Doane University, Crete, Nebraska, USA

ABSTRACT Three mycobacteriophages, Bipolarisk, Bread, and FudgeTart, were isolated from enriched soil samples found in Crete, NE. All three phages are lytic, belong to subcluster C1, and infect *Mycobacterium smegmatis* mc²155. The structures of the three genomes are similar, with slight variations in gene number and content.

Bacteriophages have the potential to be used therapeutically against bacterial pathogens (1). Mycobacteriophages are viruses that specifically infect and lyse mycobacterial hosts (2). Three phages, Bipolarisk, Bread, and FudgeTart, were isolated from soil samples collected on the Doane University campus in Crete, NE. Samples were enriched with *Mycobacterium smegmatis* mc²155 in 7H9 broth at 37°C. Phages were purified and amplified by growth on 7H9 medium with the host. All three are lytic and classified as subcluster C1 mycobacteriophages. Each phage is a myovirus, a trait shared with other cluster C phages (2). Genomic DNA was extracted from lysates using the Wizard DNA extraction kit (Promega).

DNA was sequenced at The Pittsburgh Bacteriophage Institute. Sequencing libraries were prepared from genomic DNA using an NEB Ultra II kit with dual-indexed barcoding. Libraries were pooled and run on an Illumina MiSeq instrument, yielding at least 100,000 single-end 150-base pair reads and at least 600-fold coverage for each genome. No further quality control or adapter trimming was performed on the reads provided by the sequencer. These raw reads were assembled using Newbler v.2.9 (default settings). Each genome yielded a single contig, which was checked for completeness, accuracy, and phage genomic termini using Consed v.29 as previously described (3). Gene start sites were predicted using GLIMMER v.3.02 (4) and GeneMark v.2.8 (5) using default settings and manually checked for appropriate gaps/overlaps, ribosome binding site (RBS) scores, and BLAST alignments to related phages using DNA Master v.5.23.2. (<http://cobamide2.bio.pitt.edu/computer.htm>). tRNAs were identified using ARAGORN v.1.1 (6) and tRNAscan-SE v.2.0 (6) using default settings. Predicted gene functions were assigned using NCBI BLAST (7), PECAAN (8), PhagesDB.org (9), Starterator v.1.1 (<https://seaphages.org/software/>), HHPRED v.2.0.16 (10), and Phamerator (11).

The genome lengths for Bipolarisk, Bread, and FudgeTart range from 153,796 bp to 154,734 bp, with G+C contents ranging from 64.7% to 64.8% (Table 1), similar to the host bacterium (5). All three phages have a similar number of genes with predicted functions and tRNAs (Table 1), with each containing one transfer-messenger RNA (tmRNA). The phage genomes are highly similar, with greater than 99.4% nucleotide identity and coverage of 96% to 97%, determined by nucleotide BLAST.

The genomes of Bipolarisk, Bread, and FudgeTart are circularly permuted and canonically arranged (9), with the majority of binding and attachment genes collocated near the 5' end of the annotated sequence, structural and viral assembly genes in the middle, and lysis cassette genes near the 3' end. Despite structural similarities, there are a few notable differences in specific genes between the three genomes, such as

Citation Bowder DM, Gannon BW, Grint KJ, Iltz JT, Koch TM, Mahnke KS, Murnan BD, Osborn AM, Schreiber DM, Su GL, Troester JG, Doyle EL. 2019. Genome sequences of three cluster C mycobacteriophages, Bipolarisk, Bread, and FudgeTart. *Microbiol Resour Announc* 8:e00290-19. <https://doi.org/10.1128/MRA.00290-19>.

Editor John J. Dennehy, Queens College

Copyright © 2019 Bowder et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Erin L. Doyle, erin.doyle@doane.edu.

Received 20 March 2019

Accepted 19 June 2019

Published 11 July 2019

TABLE 1 Characteristics of Bipolarisk, Bread, and FudgeTart

| Phage name | GenBank accession no. | SRA accession no. | Approx shotgun coverage (×) (no. of reads) | Genome size (bp) | No. of genes | No. of genes with predicted functions | G+C content (%) | No. of tRNAs |
|------------|--------------------------|----------------------------|--|------------------|--------------|---------------------------------------|-----------------|--------------|
| Bipolarisk | MK450429 | SRX4877857 | 653 (743,920) | 154,734 | 231 | 58 | 64.7 | 33 |
| Bread | MH779498 | SRX4877856 | 600 (653,424) | 153,796 | 230 | 60 | 64.8 | 31 |
| FudgeTart | MH779502 | SRX4877858 | 621 (680,367) | 154,658 | 231 | 61 | 64.8 | 31 |

FudgeTart_104, which differs from Bipolarisk_104 and Bread_105. The amino acid sequence for this gene in Bipolarisk and Bread is 54% identical to the gene in FudgeTart at the same position. This gene in FudgeTart is rarely called in cluster C phages and is present in all currently known cluster DO phages, which infect *Gordonia terrae*. Additionally, FudgeTart does not contain the gene directly upstream (Bipolarisk_103 and Bread_104).

Data availability. Genome sequences and raw sequence data are available in GenBank and the Sequence Read Archive. Accession numbers are shown in Table 1.

ACKNOWLEDGMENTS

We thank The Howard Hughes Medical Institute SEA-PHAGES program, The University of Pittsburgh Sequencing Center for helping us sequence the genomes, Alexander Plummer, Nick Scalora, Olivia DeFord, and Syerra Watson for discovering mycobacteriophages Bipolarisk, Bread, and FudgeTart, and Gage Herron for help with annotation.

REFERENCES

- Young R, Gill JJ. 2015. Phage therapy redux—what is to be done? *Science* 350:1163–1164. <https://doi.org/10.1126/science.aad6791>.
- Hatfull GF. 2014. Molecular genetics of mycobacteriophages. *Microbiol Spectr* 2:1–36. <https://doi.org/10.1128/microbiolspec.MGM2-0032-2013>.
- 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.
- 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>.
- 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>.
- 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>.
- Brown GR, Hem V, Katz KS, Ovetsky M, Wallin C, Ermolaeva O, Tolstoy I, Tatusova T, Pruitt KD, Maglott DR, Murphy TD. 2015. Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res* 43:D36–D42. <https://doi.org/10.1093/nar/gku1055>.
- Ting YS, Egertson JD, Bollinger JG, Searle BC, Payne SH, Noble WS, MacCoss MJ. 2017. PECAAN: library free peptide detection for data-independent acquisition tandem mass spectrometry data. *Nat Methods* 14:903–908. <https://doi.org/10.1038/nmeth.4390>.
- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.
- 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>.
- 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>.