Complete Genome Sequences of Nine *Bacillus cereus* Group Phages

Samantha Foltz, (iD) Allison A. Johnson, on behalf of the 2013–2015 VCU Phage Hunters

Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, Virginia, USA

We report the sequences of nine novel *Bacillus cereus* group bacteriophages: DIGNKC, Juglone, Nemo, Nigalana, NotTheCreek, Phrodo, SageFayge, Vinny, and Zuko. These bacteriophages are double-stranded DNA-containing *Myoviridae* isolated from soil samples using *B. thuringiensis* subsp. *kurstaki* as the host bacterium.

Address correspondence to Allison A. Johnson, aajohnson@vcu.edu.

The *Bacillus cereus* group includes 11 closely related species of rod-shaped, sporulating, Gram-positive bacteria (1). The "ATC family" (*B. anthracis*, *B. thuringiensis*, *B. cereus*) are pathogens: *B. anthracis* is an obligate human pathogen, *B. cereus* is an opportunistic pathogen that can cause food poisoning, and *B. thuringiensis* is an insect pathogen used worldwide as a biopesticide. Students at Virginia Commonwealth University, as part of the SEA PHAGES program (2), have discovered, characterized, and annotated the genomes of nine phages infecting *B. thuringiensis* subsp. *kurstaki*. Phages were isolated from soil samples collected across the United States, through the enrichment plating method, and purified by multiple rounds of plaque picking, serial dilution, and plating into *B. thuringiensis* and grown at 30°C.

The reported bacteriophages are *Myoviridae* morphology with head diameters of ~80 nm and contractile tails of ~200 nm in length. The genomic DNA of these bacteriophages was sequenced by 454 or MiSeq next generation sequencing technology to at least 100-fold coverage, assembled by Newbler software, and visualized by Consed. The physical genome ends containing long terminal repeats were determined by identifying a region with approximately double read coverage. DNAMaster (http://cobamide2.bio.pitt.edu /computer.htm) was used to annotate the genes in each *Bacillus* phage genome, and the tool integrates GeneMark (3) and Glimmer (4) for predicting open reading frames and Aragon (5) to predict the presence of tRNA genes. Functional predictions were derived from BLASTp (6) and HHpred (7) and comparing each gene to homologs in Phamerator (8) comparative genomics software.

Sequencing revealed double-stranded DNA genomes of 159,336 to 164,443 bp in length, with G+C content of 37.8 to 38.7%, and long terminal repeat sequences of 2,585 to 3,171 bp in length. Annotation predicted 288 to 302 protein coding genes for each genome. Three of the phages have three to seven tRNA genes.

Nucleotide sequence accession numbers. The complete genome sequences of the *Bacillus* phages are available in GenBank at the following accession numbers: DIGNKC, KU737349; Juglone, KU737345; Nemo, KU737352; Nigalana, KU737344; NotTheCreek, KU737351; Phrodo, KU737347; SageFayge, KU737350; Vinny, KU737346; Zuko, KU737348.

REFERENCES

1. **Liu Y, Lai Q, Göker M, Meier-Kolthoff JP, Wang M, Sun Y, Wang L, Shao Z.** 2015. Genomic insights into the taxonomic status of the *Bacillus cereus* group. Sci Rep **5:**14082. http://dx.doi.org/10.1038/srep14082.
2. **Jordan TC, Burnett SH, Carson S, Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, 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.
3. **Besemer J, Borodovsky M.** 2005. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res **33:** W451–W454. http://dx.doi.org/10.1093/nar/gki487.
4. **Salzberg SL, Delcher AL, Kasif S, White O.** 1998. Microbial gene identification using interpolated Markov models. Nucleic Acids Res **26:**544–548. http://dx.doi.org/10.1093/nar/26.2.544.
5. **Laslett D, Canback B.** 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res **32:**11–16. http://dx.doi.org/10.1093/nar/gkh152.
6. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389–3402. http://dx.doi.org/10.1093/nar/25.17.3389.
7. **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.
8. **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.