AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of Two Temperate *Bacillus subtilis* Phages Isolated at Tumamoc Hill Desert Laboratory

Greg P. Krukonis,[a] Amanda K. Kemp,[a] Katie F. Storrie,[a] Vivian R. Chavira,[a] Hayden W. Lantrip,[a] Victoria D. Perez,[a] Desiree A. Reyes,[a] Julian A. Truax,[a] Rachel Loney,[b] Véronique A. Delesalle[b]

[a]Department of Biology, Angelo State University, San Angelo, Texas, USA
[b]Department of Biology, Gettysburg College, Gettysburg, Pennsylvania, USA

ABSTRACT Bacteriophages are important in structuring bacterial communities, including desert soils dominated by *Bacillus* species. Here, we describe two genetically similar temperate phages isolated on a *Bacillus subtilis* strain from soil in Tucson, Arizona. Their double-stranded DNA (dsDNA) genomes contain 98 and 102 genes, with a set of 4 genes being found in only one phage.

Bacterial communities in desert environments are often dominated by *Firmicutes* strains, including *Bacillus subtilis* and relatives (1–4). Given how bacteriophages impact bacterial communities (5–7), understanding these communities requires understanding phage diversity. Here, we describe two temperate phages from the Sonoran Desert.

Each phage was isolated from its own soil sample collected at Tumamoc Hill Desert Laboratory (Tucson, AZ) in May 2016 (32°13′04.9″N, 111°00′12.9″W), at sites separated by 10 m. The soil was dry and sandy, dug to 10 cm. Approximately 1 g of soil was added to 20 mL LB broth, incubated for 4 h at 37°C with shaking at 250 rpm, and then filtered (0.22 μm). Samples were then plated on *Bacillus subtilis* strain T89-06 (also called S89-6 or T89-6), which was originally isolated by Istock and colleagues (8, 9). Individual plaques were isolated and were single plaque purified three times on lawns made from spores of the isolation host. High-titer lysates were prepared by flooding, with LB broth, multiple plates containing at least 10^4 plaques. Lysates were filtered, and DNA was extracted using phenol-chloroform (10). For sequencing, libraries were prepared with the Illumina TruSeq Nano DNA library preparation kit and sequenced with the Illumina MiSeq platform, using a 150-bp single-end read v3 flow cell, at the North Carolina State University Genomic Science Laboratory. We assembled genomes using GS De Novo Assembler v2.9 (11). For each phage, the 150-bp reads were assembled into one contig with >1,000× coverage, and contig consensus quality was verified in Consed v29 (12) (Table 1). Genome ends were determined with PhageTerm (13) (Table 1). The finished sequences were imported into DNA Master v5.22.22 (http://cobamide2.bio.pitt.edu/computer.htm) to map and compare open reading frames. Putative genes were called based on both Glimmer v3.0 and GeneMark v2.5 algorithms (14, 15). Putative protein functions were predicted using BLAST v2.12 (16) and HHpred (17). For BLASTp matches, an E value of $<10^{-5}$ was required to assign function. For HHpred matches, a high probability (>85%), substantial coverage (>50%), and low E value ($<10^{-5}$) were required. The absence of tRNA genes was confirmed with ARAGORN (18). Default settings were used for all programs.

Phages 268TH002 and 268TH007 have double-stranded DNA (dsDNA) genomes with 98 and 102 predicted protein coding genes, respectively (Table 1), and a genome organization typical of *Siphoviridae*, with structural genes showing conserved order (19). They show limited nucleotide similarity to other sequenced phages (Table 1) but

Editor Simon Roux, DOE Joint Genome Institute

Address correspondence to Greg P. Krukonis, gkrukonis@angelo.edu.

The authors declare no conflict of interest.

Received 10 July 2022
Accepted 30 July 2022
Published 17 August 2022

TABLE 1 Sequencing information and genome characteristics for *Bacillus* phages 268TH002 and 268TH007

Phage name	No. of reads	Coverage (×)	Genome size (bp)	GC content (%)	Genome ends[a]	No. of protein-coding genes	Best BLASTn match (GenBank accession no.)[b]	Query coverage (%) with best match	Identity (%) with best match
268TH002	807,065	1,868	65,534	47.3	310-bp DTRs	98	*Bacillus velezensis* strain Lzh-a42 (CP025308.1)	63	88
							Bacillus phage vB_BauS_KLEB27-1 (OM654379.1)	24	76
268TH007	505,753	1,124	68,062	47.4	310-bp DTRs	102	*Bacillus velezensis* strain Lzh-a42 (CP025308.1)	61	88
							Bacillus phage vB_BauS_KLEB27-1 (OM654379.1)	23	76

[a] DTR, direct terminal repeat. By convention, genomes start and end with the DTR sequence and with the terminase gene on the forward strand (11).
[b] The genome of each phage was compared to the complete nucleotide database and to the same database restricted to all tailed phages (combined taxid numbers 10699, 10662, and 10744) with BLASTn. For each search, the best match is reported.

share 96% nucleotide identity with each other, differing primarily through the presence of 2,525 bp in the middle of the genome of 268TH007 with four open reading frames (putatively coding for a FtsK-like DNA translocase, a replication-relaxation family protein, a helix-turn-helix transcriptional regulator, and a hypothetical protein). FtsK translocases are involved in the bacterial SOS response to DNA damage, can activate prophage induction (20), and may broaden conditions for prophage induction. In addition, two genes whose predicted products have sequence identity to tyrosine recombinase have been identified. Whether and how these function in phage integration are open questions. Finally, both phages have ribonucleotide reductase genes, which may benefit them through synthesis of deoxyribonucleotides during periods when host DNA synthesis is inactive (21).

Data availability. Genome sequences and associated information can be found under the following GenBank and SRA accession numbers: 268TH002, ON210835 and SRX15148566; 268TH007, ON210834 and SRX15148567, respectively.

ACKNOWLEDGMENTS

We thank Angelo State University and Gettysburg College (through research and professional development grants to V.A.D.) for financial support of this research.

REFERENCES

1. Roberts MS, Cohan FM. 1995. Recombination and migration rates in natural populations of *Bacillus subtilis* and *Bacillus mojavensis*. Evolution 49:1081–1094. https://doi.org/10.1111/j.1558-5646.1995.tb04435.x.
2. Chanal A, Chapon V, Benzerara K, Barakat M, Christen R, Achouak W, Barras F, Heulin T. 2006. The desert of Tataouine: an extreme environment that hosts a wide diversity of microorganisms and radiotolerant bacteria. Environ Microbiol 8:514–525. https://doi.org/10.1111/j.1462-2920.2005.00921.x.
3. Prestel E, Salamitou S, DuBow MS. 2008. An examination of the bacteriophages and bacteria of the Namib desert. J Microbiol 46:364–372. https://doi.org/10.1007/s12275-008-0007-4.
4. Prestel E, Regeard C, Salamitou S, Neveu J, DuBow M. 2013. The bacteria and bacteriophages from a Mesquite Flats site of the Death Valley desert. Antonie Van Leeuwenhoek 103:1329–1341. https://doi.org/10.1007/s10482-013-9914-4.
5. Koskella B, Meaden S. 2013. Understanding bacteriophage specificity in natural microbial communities. Viruses 5:806–823. https://doi.org/10.3390/v5030806.
6. Fuhrman JA, Schwalbach M. 2003. Viral influence on aquatic bacterial communities. Biol Bull 204:192–195. https://doi.org/10.2307/1543557.
7. Thurber RV. 2009. Current insights into phage biodiversity and biogeography. Curr Opin Microbiol 12:582–587. https://doi.org/10.1016/j.mib.2009.08.008.
8. Duncan KE, Ferguson N, Kimura K, Zhou X, Istock C. 1994. Fine-scale genetic and phenotypic structure in natural populations of *Bacillus subtilis* and *Bacillus licheniformis*: implications for bacterial evolution and speciation. Evolution 48:2002–2025. https://doi.org/10.1111/j.1558-5646.1994.tb02229.x.
9. Istock CA, Ferguson N, Istock NL, Duncan KE. 2001. Geographical diversity of genomic lineages in *Bacillus subtilis*. Org Divers Evol 1:179–191. https://doi.org/10.1078/1439-6092-00017.
10. Green MR, Sambrook J. 2017. Protocol 12: isolation of high-molecular-weight DNA using organic solvents. Cold Spring Harb Protoc 2017:44–46. https://doi.org/10.1101/pdb.prot093450.
11. Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes. Methods Mol Biol 1681:109–135. https://doi.org/10.1007/978-1-4939-7343-9_9.
12. Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. Bioinformatics 29:2936–2937. https://doi.org/10.1093/bioinformatics/btt515.
13. Garneau JR, Depardieu F, Fortier LC, Bikard D, Monot M. 2017. PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. Sci Rep 7:8292. https://doi.org/10.1038/s41598-017-07910-5.
14. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27:4636–4641. https://doi.org/10.1093/nar/27.23.4636.
15. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res 26:1107–1115. https://doi.org/10.1093/nar/26.4.1107.
16. Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL. 2012. Domain enhanced lookup time accelerated BLAST. Biol Direct 7:12. https://doi.org/10.1186/1745-6150-7-12.
17. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:244–248. https://doi.org/10.1093/nar/gki408.

18. 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.
19. Hatfull GF. 2012. The secret lives of mycobacteriophages. Adv Virus Res 82:179–288. https://doi.org/10.1016/B978-0-12-394621-8.00015-7.
20. Wolfe A, Phipps K, Weitao T. 2014. Viral and cellular SOS-regulated motor proteins: dsDNA translocation mechanisms with divergent functions. Cell Biosci 4:31. https://doi.org/10.1186/2045-3701-4-31.
21. Dwivedi B, Xue B, Lundin D, Edwards RA, Breitbart M. 2013. A bioinformatic analysis of ribonucleotide reductase genes in phage genomes and metagenomes. BMC Evol Biol 13:33. https://doi.org/10.1186/1471-2148-13-33.