AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of Cluster C2 *Bacillus* Phage Maceta

Elizabeth Greguske[a]

[a]Department of Biology, Saint Anselm College, Manchester, New Hampshire, USA

ABSTRACT *Bacillus* phage Maceta was isolated from the soil of commercially purchased annual flowers using the host *Bacillus thuringiensis* serovar kurstaki. Isolated DNA was then sequenced and annotated. Maceta has a relatively small genome, containing 45,023 bp, and shares an average nucleotide identity of 96% with other cluster C2 *Bacillus* phage.

Maceta is a novel *Bacillus* bacteriophage isolated, characterized, and annotated by an undergraduate student research lab associated with the Howard Hughes Medical Institute (HHMI) Phage Hunters (SEA-PHAGES) program (1). Maceta, named for the flower pot from which the soil was sampled, was isolated in 2014 using *Bacillus thuringiensis* serovar kurstaki as the host bacterium. Both the soil and the annual flowers in the outdoor pot (Saint Anselm College, Manchester, NH, USA) were commercially purchased and then planted by groundskeepers 2 weeks prior to sampling. The soil sample was added to 45 ml of Trypticase soy broth (TSB) and 5 ml of log-phase *Bacillus thuringiensis* serovar kurstaki, mixed well, and then grown overnight at 37°C. The supernatant was collected and centrifuged at 3,000 rpm for 10 min to pellet the remaining soil and *Bacillus thuringiensis* serovar kurstaki. The pellet was discarded, and the remaining liquid was sterile filtered through a 22-μm filter with vacuum suction (2). The resulting filtrate was tested for sufficient phage concentration ($>10^8$ PFU/ml) via serial dilution and plate counts. DNA was isolated from purified phage using a Promega Wizard DNA kit and sequenced via the MiSeq Illumina platform at the Hubbard Center for Genome Studies (University of New Hampshire [UNH], Durham, NH, USA), resulting in 9,530,240 paired reads, with a read length average of 151 bp. Trimmomatic v0.33 with default settings was used to trim the reads (3). Geneious v10.2 with default settings (4) was used to assemble the genome, and the quality of the assembly was assessed by Quast v4.0 (5). The genome was then further refined using Geneious v10.2 with default settings except for sensitivity, which was set to highest sensitivity/slow. The average depth of coverage was 3,984.2×, with no areas of poor coverage noted. Maceta was autoannotated in Geneious v10.2 with default settings, using *Bacillus* phage Bastille, GenBank accession number JF966203 (6), for comparison. Maceta was then manually cross-checked against 4 other closely related cluster C *Bacillus* phage genomes, namely CAM003, Evoli, Vinny, and Anthony, registered under GenBank accession numbers KJ489397 (7), KJ489398 (7), KU737346 (8), and MF498901 (9), respectively.

Maceta has a relatively short genome compared with those of other *Bacillus* phage (10), containing 45,023 bp with a GC content of 40.0%. The genome contains 46 predicted genes, 37 of which are protein-coding regions and 9 of which are tRNA-coding regions. Of the 37 protein-coding regions, 19 were assigned a function, all of which are involved in capsid and tail production and assembly or host cell attachment and penetration. No endonucleases were identified, strongly suggesting that Maceta has an obligatory lytic life cycle and is unable to undergo the lysogenic life cycle. Maceta contains 9 tRNAs, 8 of which are clustered together near the end of the linear

Received 20 September 2018 **Accepted** 15 November 2018 **Published** 13 December 2018

Citation Greguske E. 2018. Complete genome sequence of cluster C2 *Bacillus* phage Maceta. Microbiol Resour Announc 7:e01298-18. https://doi.org/10.1128/MRA.01298-18.

Editor Frank J. Stewart, Georgia Institute of Technology

Address correspondence to egreguske@anselm.edu.

genome. The remaining tRNA is sandwiched between a tail fiber hydrolase protein-coding sequence and an HK97 family portal protein-coding sequence.

Cluster assignment was based on average nucleotide identity (ANI) as previously described (7, 10). Maceta had a whole-genome ANI of 96% with 3 C2 phage (6, 9), as determined by BLASTn (11), firmly placing Maceta in this widespread cluster.

Data availability. The complete genome sequence of *Bacillus* phage Maceta is available in GenBank under accession number MH538296. Raw sequence reads are available in SRA under BioProject accession number PRJNA498083. Maceta is also registered with the HHMI Phage Hunters Program at bacillus.phagesdb.org (12).

ACKNOWLEDGMENT

This research was supported by New Hampshire-INBRE through an Institutional Development Award (IDeA), number P20GM103506, from the National Institute of General Medical Sciences of the NIH.

REFERENCES

1. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SC, 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. https://doi.org/10.1128/mBio.01051-13.
2. Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science. 2011. Bacillus phage laboratory manual. Howard Hughes Medical Institute, Chevy Chase, MD.
3. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
4. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
5. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
6. Jarvis AW, Collins LJ, Ackermann HW. 1993. A study of five bacteriophages of the Myoviridae family which replicate on different Gram-positive bacteria. Arch Virol 133:75–84. https://doi.org/10.1007/BF01309745.
7. Sauder AB, Quinn MR, Brouillette A, Caruso S, Cresawn S, Erill I, Lewis L, Loesser-Casey K, Pate M, Scott C, Stockwell S, Temple L. 2016. Genomic characterization and comparison of seven Myoviridae bacteriophage infecting Bacillus thuringiensis. Virology 489:243–251. https://doi.org/10.1016/j.virol.2015.12.012.
8. Foltz S, Johnson AA. 2016. Complete genome sequences of nine Bacillus cereus group phages. Genome Announc 4:e00473-16. https://doi.org/10.1128/genomeA.00473-16.
9. Lee M, Puglisi KM, Cohort US-B, Erill I, Caruso SM. 2018. Complete genome sequences of HonestAbe, Anthony, and Taffo16, three cluster C Bacillus cereus group bacteriophages. Genome Announc 6:e00493-18. https://doi.org/10.1128/genomeA.00493-18.
10. Grose JH, Jensen GL, Burnett SH, Breakwell DP. 2014. Genomic comparison of 93 Bacillus phages reveals 12 clusters, 14 singletons and remarkable diversity. BMC Genomics 15:855. https://doi.org/10.1186/1471-2164-15-855.
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/S0022-2836(05)80360-2.
12. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711.