



Complete Genome Sequences of *Bacillus* Phages DirtyBetty and Kida

Kelly Flounlacker, Rachel Miller, Diana Marquez,  Allison Johnson, on behalf of the 2015–2016 VCU Phage Hunters

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

ABSTRACT We report here the genome sequences of two bacteriophages of the *Bacillus cereus* group, DirtyBetty and Kida. These bacteriophages are double-stranded DNA-containing *Myoviridae* isolated from soil samples using *Bacillus thuringiensis* subsp. *kurstaki* as their host bacteria.

The ATC family of the genus *Bacillus* includes the spore-forming soil bacteria *B. cereus*, *B. anthracis*, and *B. thuringiensis*. The bacteria in the ATC family are very closely related genomically but display pathogenic differences (1, 2). It is of interest to discover phages which can be isolated safely by students using *B. thuringiensis* to learn more about these phages and their host bacteria in the ATC group. Students in the VCU Phage Lab, as part of the Howard Hughes Medical Institute (HHMI) SEA-PHAGES program (3), have isolated bacteriophage infecting *Bacillus* spp. for the past 3 years. These phages are generally *Myoviridae* with isometric heads containing double-stranded DNA and long contractile tails.

Phages Kida and DirtyBetty were isolated from soil samples collected in Richmond and Vienna, VA, respectively, using enrichment with *B. thuringiensis* subsp. *kurstaki* as the host bacterium. Phages Kida and DirtyBetty both have head diameters of 100 nm, as well as tail lengths of 200 nm and 375 nm, respectively. Kida forms clear plaques on a lawn of *B. thuringiensis*, while DirtyBetty forms plaques with thin halos. The genomic DNA of these bacteriophages was sequenced by MiSeq next-generation sequencing technology to at least 45-fold coverage, assembled by the 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. DNA Master (<http://cobamide2.bio.pitt.edu/computer.htm>) was used to annotate the genes in each genome, and the software tool integrates GeneMark (4) and Glimmer (5) for predicting open reading frames and Aragorn (6) to predict the presence of tRNA genes. Functional predictions were derived from BLASTP (7) conserved domains and HHpred (8) and by comparing each gene to homologs in the Phamerator (9) comparative genomics software.

The Kida genome has a length of 162,151 bp, with long terminal repeats at a length of 2,574 bp. DirtyBetty has a genome length of 162,415, with 2,808-bp long terminal repeats. Both genomes have 38.7% G+C content and have 300 (DirtyBetty) and 304 (Kida) predicted open reading frames. Neither phage has tRNAs present in its sequence. The two phages share 94% query coverage and 95% identity to each other, as well as similarly high values to other published cluster C1 *Bacillus* phages (10–13). They share 278 protein families (92% of their proteins), or phams (as described in reference 9), and relatively few protein-coding genes that are present in only one of these phages. Furthermore, only 15% of these proteins (44/300 for DirtyBetty and 43/304 for Kida) have a predicted function, generally for structural, lytic, and nucleotide metabolism genes. There are no genes in these genomes without at least one homolog in our *Bacillus* phage database (<http://bacillus.phagesdb.org/>), which indicates that no novel genes are present in these two phages. While the genes with predicted function are

Received 14 October 2016 Accepted 11 January 2017 Published 9 March 2017

Citation Flounlacker K, Miller R, Marquez D, Johnson A, 2015–2016 VCU Phage Hunters. 2017. Complete genome sequences of *Bacillus* phages DirtyBetty and Kida. *Genome Announc* 5:e01385-16. <https://doi.org/10.1128/genomeA.01385-16>.

Copyright © 2017 Flounlacker 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 Allison Johnson, aajohnson@vcu.edu.

conserved in these phages and other closely related cluster C1 phages, this group also contains a large set of conserved proteins with no identifiable homologs outside this group, providing a rich data set for determination of gene essentiality and function.

Accession number(s). The complete genome sequences of the *Bacillus* phages DirtyBetty and Kida are available in GenBank under accession numbers [KX349903](#) and [KX349902](#), respectively.

ACKNOWLEDGMENTS

This work was supported by VCU Life Sciences, the VCU Nucleic Acids Research Facility, and the HHMI SEA-PHAGES program. The members of the 2015–2016 VCU Phage Hunters courses are listed at <https://wiki.vcu.edu/display/phagelab/VCU+Phage+Hunters>.

REFERENCES

1. Yang L, Qiliang L, Markus G, Meier-Kolthoff JP, Meng W, Yamin S, Lei W, Zongze S. 2015. Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci Rep* 5:14082.
2. El-Arabi TF, Griffiths MW, She YM, Villegas A, Lingohr EJ, Kropinski AM. 2013. Genome sequence and analysis of a broad-host range lytic bacteriophage that infects the *Bacillus cereus* group. *Virol J* 10:48. <https://doi.org/10.1186/1743-422X-10-48>.
3. 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>.
4. 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>.
5. Salzberg SL, Delcher AL, Kasif S, White O. 1998. Microbial gene identification using interpolated markov models. *Nucleic Acids Res* 26:544–548. <https://doi.org/10.1093/nar/26.2.544>.
6. 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>.
7. 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. <https://doi.org/10.1093/nar/25.17.3389>.
8. 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>.
9. 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>.
10. 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>.
11. Erill I, Caruso SM. 2015. Genome sequences of two *Bacillus cereus* group bacteriophages, Eyuki and AvesoBmore. *Genome Announc* 3(5):e01199-15. <https://doi.org/10.1128/genomeA.01199-15>.
12. Foltz S, Johnson AA. 2013–2015 VCU Phage Hunters. 2016. Complete genome sequences of nine *Bacillus cereus* group phages. *Genome Announc* 4(4):e00473-16. <https://doi.org/10.1128/genomeA.00473-16>.
13. Cornell JL, Breslin E, Schuhmacher Z, Himelright M, Berluti C, Boyd C, Carson R, Del Gallo E, Giessler C, Gilliam B, Heatherly C, Nevin J, Nguyen B, Nguyen J, Parada J, Sutterfield B, Tukruni M, Temple L. 2016. Complete genome sequence of *Bacillus thuringiensis* bacteriophage smudge. *Genome Announc* 4(4):e00572-16. <https://doi.org/10.1128/genomeA.00572-16>.