



# Complete Genome Sequence of *Bacillus cereus* Bacteriophage vB\_BceS\_KLEB30-3S

Monika Šimoliūnienė,<sup>a</sup> Deividas Tumėnas,<sup>a</sup> Kotryna Kvederavičiūtė,<sup>b</sup> Rolandas Meškys,<sup>a</sup> Sigitas Šulčius,<sup>c</sup>  
Eugenijus Šimoliūnas<sup>a</sup>

<sup>a</sup>Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Vilnius, Lithuania

<sup>b</sup>Department of Biological DNA Modification, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania

<sup>c</sup>Laboratory of Algology and Microbial Ecology, Nature Research Centre, Vilnius, Lithuania

**ABSTRACT** In this study, we present the genomic characterization of the temperate bacteriophage vB\_BceS\_KLEB30-3S (KLEB30-3S), which was induced from *Bacillus cereus* strain KR3M-30, isolated from a gypsum karst lake ecosystem in Lithuania. The 37,134-bp genome of KLEB30-3S contains 58 predicted protein-encoding genes and no tRNA genes.

*Bacillus cereus* comprises a highly versatile group of bacteria which are of special interest because of their ecological significance, broad range of pathogenicity, and biocontrol potential (1, 2). A number of *Bacillus* phages and their components could be very useful tools for typing or detecting their hosts, significantly contribute to bacterial genetic diversity, and provide the host with special characteristics such as pathogenicity and distinct phenotypic traits (3–5). Thus, new isolation and insights into *Bacillus* phages remain highly relevant.

Here, we report the complete genome sequence of a temperate bacteriophage, vB\_BceS\_KLEB30-3S (KLEB30-3S), induced from *Bacillus cereus* strain KR3M-30 when cultured with KR3M-30-specific bacteriophage vB\_BceS\_KLEB30 (our unpublished data). The host strain was derived from a gypsum karst lake ecosystem in Biržai, Lithuania (56°14'55.5"N, 24°41'33.7"E), cultivated aerobically in LB at 22°C and identified using 16S rRNA gene sequence analysis. Phage KLEB30-3S was propagated using the soft-agar overlay method and purified using CsCl gradient ultracentrifugation as described previously (6). Phage DNA was isolated using the phenol-chloroform extraction and ethanol precipitation method (7). The complete genome sequence of KLEB30-3S was determined using Illumina DNA sequencing technology at BaseClear (Leiden, The Netherlands). Briefly, the sequencing library was prepared using in-house-developed library preparation solutions (BaseClear). Paired-end sequence reads were generated using the Illumina NovaSeq 6000 system (Illumina, San Diego, CA, USA). FASTQ read sequence files were generated using bcl2fastq2 version 2.18. Initial quality assessment was based on data passing the Illumina chastity filtering. Subsequently, reads containing the PhiX control signal were removed using an in-house filtering protocol (BaseClear). In addition, reads containing (partial) adapters were clipped (to a minimum read length of 50 bp). The second quality assessment was based on the remaining reads using the FastQC quality control tool version 0.11.5. The quality of the Illumina reads was improved using the error correction tool BayesHammer (8). Error-corrected reads were assembled into contigs using SPAdes version 3.10 (9). The contigs were linked together and placed into scaffolds using SSPACE version 2.3 (10). Using Illumina reads, gapped regions within scaffolds were (partially) closed using GapFiller version 1.10 (11). Finally, assembly errors and the nucleotide disagreements between the Illumina reads and scaffold sequences were corrected using Pilon version 1.21 (12).

**Citation** Šimoliūnienė M, Tumėnas D, Kvederavičiūtė K, Meškys R, Šulčius S, Šimoliūnas E. 2020. Complete genome sequence of *Bacillus cereus* bacteriophage vB\_BceS\_KLEB30-3S. Microbiol Resour Announc 9:e00348-20. <https://doi.org/10.1128/MRA.00348-20>.

**Editor** Simon Roux, DOE Joint Genome Institute

**Copyright** © 2020 Šimoliūnienė 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 Eugenijus Šimoliūnas, [eugenijus.simoliunas@bchi.vu.lt](mailto:eugenijus.simoliunas@bchi.vu.lt).

**Received** 2 April 2020

**Accepted** 12 April 2020

**Published** 14 May 2020

**TABLE 1** Genomic characteristics of *Bacillus* phage KLEB30-3S and its closest relatives

| Phage             | GenBank accession no.     | Genome size (bp) | G+C content (%) | No. of ORFs | Coverage with KLEB30-3S (%) | Identity with KLEB30-3S (%) |
|-------------------|---------------------------|------------------|-----------------|-------------|-----------------------------|-----------------------------|
| vB_BceS_KLEB30-3S | <a href="#">MT136606</a>  | 37,134           | 38.3            | 58          |                             |                             |
| vB_BtS_BMBtp2     | <a href="#">NC_019912</a> | 36,932           | 37.8            | 53          | 87                          | 91.30                       |
| proCM3            | <a href="#">KF_296717</a> | 43,278           | 37.4            | 66          | 87                          | 90.56                       |
| TP21-L            | <a href="#">NC_011645</a> | 37,456           | 37.8            | 56          | 83                          | 90.19                       |
| BMBtpLA3          | <a href="#">KX_190834</a> | 37,385           | 37.8            | 59          | 83                          | 90.19                       |

In total, 168,858 high-quality reads were obtained and further assembled into one contig of 37,134 bp with an average coverage of 1,300-fold. The physical ends of the viral genome were identified using PhageTerm (13). Open reading frames (ORFs) were predicted with Geneious Prime 2019. The analysis of the genome sequence was performed using Transeq ([http://www.ebi.ac.uk/Tools/st/emboss\\_transeq](http://www.ebi.ac.uk/Tools/st/emboss_transeq)), Fasta-Nucleotide, Fasta-Protein, BLASTP, and Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo>), as well as HHPred, HHblits, HMMER, and HHSenser (14). Also, tRNAscan-SE version 1.21 (<http://lowelab.ucsc.edu/tRNAscan-SE>) was used to search for tRNAs. The whole-genome comparative analysis using the online NCBI BLASTn (nucleotide collection [nr/nt] database; <https://blast.ncbi.nlm.nih.gov>) revealed that KLEB30-3S is closely related to the *Bacillus* bacteriophages from the genus *Lwoffvirus* (summarized in Table 1) (15–17). All bioinformatics tools described in this section were run with default parameters.

The genome of KLEB30-3S was predicted to be a linear, double-stranded, terminally redundant DNA molecule consisting of 37,134 bp. It has a G+C content of 38.3%, which is similar to that of *Bacillus cereus* (34.8 to 35.5%) (18). The genome of KLEB30-3S is closely packed with an average ORF size of 592 bp, and 91.82% of the genome is coding. It has 58 probable protein-encoding genes and no genes for tRNAs. While most of the KLEB30-3S genes were found to initiate from AUG (51 out of 58 ORFs), 4 ORFs were found to initiate with GUG and 3 with UUG. A marked asymmetry in the distribution of the genes on the two phage KLEB30-3S DNA strands was observed. The vast majority (56 out of 58) of KLEB30-3S ORFs have been predicted to be transcribed from the same DNA strand, whereas only two ORFs have been found on the opposite strand. Based on homology to biologically defined proteins, 19 ORFs of KLEB30-3S were given a putative functional annotation. The genome of KLEB30-3S appears to have a modular organization, with genes for DNA packaging (terminase small and large subunits), structure/morphogenesis (head morphogenesis protein, portal protein, scaffold protein, major capsid protein, head-tail connector, head-tail adaptor, tape measure protein, tail protein, and tail fiber protein), host lysis (holin and endolysin), lysogeny (integrase and repressor), and DNA replication/recombination (replicative DNA helicase and transcriptional regulators) clustered together. Notably, no virulence factors or antibiotic resistance determinants were detected in the genome of KLEB30-3S.

**Data availability.** The complete genome sequence of phage KLEB30-3S is available in the GenBank database under accession number [MT136606](#). The accession number of the PCR-amplified 16S rRNA gene sequence of *Bacillus cereus* strain KR3M-30 is [MN752435](#). The raw sequence reads are available in the SRA database under accession number [SRR11441841](#) (BioProject number [PRJNA613569](#) and BioSample number [SAMN14409531](#)).

## ACKNOWLEDGMENT

This research was funded by grant number MSF-LMT-2 from Vilnius University.

## REFERENCES

1. Ceuppens S, Boon N, Uyttendaele M. 2013. Diversity of *Bacillus cereus* group strains is reflected in their broad range of pathogenicity and diverse ecological lifestyles. *FEMS Microbiol Ecol* 84:433–450. <https://doi.org/10.1111/1574-6941.12110>.
2. Shafi J, Tian H, Ji M. 2017. *Bacillus species* as versatile weapons for plant pathogens: a review. *Biotech Biotechnol Equip* 31:1–14.
3. Fujinami Y, Hirai Y, Sakai I, Yoshino M, Yasuda J. 2007. Sensitive detection of *Bacillus anthracis* using a binding protein originating from gamma-

- phage. *Microbiol Immunol* 51:163–169. <https://doi.org/10.1111/j.1348-0421.2007.tb03894.x>.
4. Gillis A, Mahillon J. 2014. Phages preying on *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*: past, present and future. *Viruses* 6:2623–2672. <https://doi.org/10.3390/v6072623>.
  5. Geng P, Tian S, Yuan Z, Hu X. 2017. Identification and genomic comparison of temperate bacteriophages derived from emetic *Bacillus cereus*. *PLoS One* 12:e0184572. <https://doi.org/10.1371/journal.pone.0184572>.
  6. Šimoliūnas E, Kaliniene L, Truncaitė L, Zajančauskaitė A, Staniulis J, Kaupinis J, Ger M, Valius M, Meškys R. 2013. *Klebsiella* phage vB\_KleM-RaK2: a giant singleton virus of the family *Myoviridae*. *PLoS One* 8:e60717. <https://doi.org/10.1371/journal.pone.0060717>.
  7. Carlson K, Miller E. 1994. Experiments in T4 genetics, p 419–483. In Karam JD (ed), *Bacteriophage T4*. ASM Press, Washington, DC.
  8. Nikolenko SI, Korobeynikov AI, Alekseyev MA. 2013. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* 14:S7. <https://doi.org/10.1186/1471-2164-14-S1-S7>.
  9. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
  10. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
  11. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. *Genome Biol* 13:R56. <https://doi.org/10.1186/gb-2012-13-6-r56>.
  12. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Ear AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
  13. Garneau JR, Depardieu F, Fortier LC, Bikard D, Monot M. 2017. Phage-Term: 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. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol* 430:2237–2243. <https://doi.org/10.1016/j.jmb.2017.12.007>.
  15. Dong Z, Peng D, Wang Y, Zhu L, Ruan L, Sun M. 2013. Complete genome sequence of *Bacillus thuringiensis* bacteriophage BMBtp2. *Genome Announc* 1:e00011-12. <https://doi.org/10.1128/genomeA.00011-12>.
  16. Klumpp J, Calendar R, Loessner MJ. 2010. Complete nucleotide sequence and molecular characterization of *Bacillus* phage TP21 and its relatedness to other phages with the same name. *Viruses* 2:961–971. <https://doi.org/10.3390/v2040961>.
  17. Yuan Y, Gao M, Wu D, Liu P, Wu Y. 2012. Genome characteristics of a novel phage from *Bacillus thuringiensis* showing high similarity with phage from *Bacillus cereus*. *PLoS One* 7:e37557. <https://doi.org/10.1371/journal.pone.0037557>.
  18. Zeng Q, Xie J, Li Y, Gao T, Xu C, Wang Q. 2018. Comparative genomic and functional analyses of four sequenced *Bacillus cereus* genomes reveal conservation of genes relevant to plant-growth-promoting traits. *Sci Rep* 8:1–10. <https://doi.org/10.1038/s41598-018-35300-y>.