



Complete genome of a cereulide-producing *Bacillus mobilis* strain

Hendrik Frentzel,¹ Ylanna Kelner-Burgos,¹ Pascal Witt,¹ Maria Borowiak,¹ Carlus Deneke¹

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT Hitherto, reported cereulide-producing strains of the *Bacillus cereus* group belonged to the *panC* clades 3 or 6 and shared the highest average nucleotide identity (ANI) with *B. paranthracis* or *B. mycoides*. Here, we report an emetic strain assigned to the *panC* clade 2 and the ANI species *B. mobilis*.

KEYWORDS *Bacillus cereus*, toxin, emetic, plasmid

The *Bacillus cereus* group contains strains that are able to produce the emetic toxin cereulide, which causes food poisoning. The genetic prerequisite for cereulide production is the *cesHPTABCD* gene cluster. Most described emetic *B. cereus* group strains are mesophilic, belong to the phylogenetic *panC* clade 3, share the highest average nucleotide identity (ANI) with the species *B. paranthracis*, and carry the *ces* genes on a 270 kb plasmid (1–3). Very rarely, psychrotolerant emetic strains are detected (*panC* clade 6 and highest ANI with *B. mycoides*) (4–6).

The emetic strain BfR-BA-01700 described in this study was isolated in accordance with ISO 7932:2004 by spread plating on mannitol egg yolk polymyxin agar, followed by incubation for 24 h to 48 h at 30°C, from a sample of meat salad in Germany in 2021 by a local food control authority. The strain belongs to the *panC* clade 2 and shares the highest ANI with *B. mobilis*. The ability of the strain to produce cereulide has been described previously (7). The *cesHPTABCD* gene cluster is located on a 349,073 bp plasmid. Using NCBI BlastN, the sequence of the classical 270 kb pCER270 plasmid (strain AH187; GenBank accession: [DQ889676.1](#)) showed a query coverage of 90% and an identity of 99.96% to the 349,073 bp plasmid.

The genomic analyses comprised short-read Illumina and long-read Oxford Nanopore Technologies sequencing. For DNA extraction, the strain was streaked from a –80°C glycerol culture on sheep blood agar. Material from a single colony was used to inoculate 3 mL of brain heart infusion broth (Oxoid), which was incubated at 30°C and 150 rpm overnight. From 1 mL of this culture, DNA was extracted with the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) and used for sequencing without further processing or size selection of the DNA.

For short-read sequencing, the Nextera DNA Flex kit (Illumina, San Diego, CA, USA) was applied, and sequencing was performed in 2×149 bp cycles on an Illumina NextSeq 500, resulting in 2,832,646 raw reads. The pipeline AQUAMIS v1.2.0-18-g1317d15 was used for trimming, quality check, and *de novo* assembly of sequencing reads (8).

For ONT long-read sequencing, the SQK-RBK110.96 kit (Oxford Nanopore Technologies, Oxford, UK) was used to prepare a sequencing library. The library was sequenced with a FLO-MIN106 flow cell and a Minlon Mk1C device. The generated data were subsequently basecalled using guppy v6.0.1 in SUP mode, resulting in 141,748 raw reads (<https://nanoporetech.com/community>). The obtained data were quality controlled and assembled using the MilongA pipeline v1.0.1 (https://gitlab.com/bfr_bioinformatics/milonga). There, reads were trimmed with Porechop v0.2.4 ([**Editor** David Rasko, University of Maryland School of Medicine, Baltimore, Maryland, USA](https://</p></div><div data-bbox=)

Address correspondence to Hendrik Frentzel, hendrik.frentzel@bfr.bund.de.

The authors declare no conflict of interest.

Received 16 July 2025

Accepted 11 November 2025

Published 24 November 2025

Copyright © 2025 Frentzel et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

TABLE 1 Genetic characteristics of the emetic *B. mobilis* strain and applied tools and databases

	Strain BfR-BA-01700	Tools and databases (all tools were used with default settings)
short-read quality control	2,755,852 reads (after trimming); 88% ≥ Q30	AQUAMIS v1.2.0-18-g1317d15 (8) utilizing fastp v0.19.5 (17)
long-read quality control	71,429 reads (after trimming and filtering); N50: 6,313 bp	Porechop v0.2.4 (https://github.com/rrwick/Porechop), NanoFilt v2.8.0 & NanoStat v1.5.0 (9)
assembly	GCA_033841795.2 (complete, seven contigs) 1. CP185965.1 : chromosome; 5,277,139 bp; circular; %GC: 35.5 2. CP185966.1 : plasmid; 349,073 bp; circular; %GC: 34 3. CP185967.1 : plasmid; 66,276 bp; circular; %GC: 32 4. CP185968.1 : plasmid; 52,852 bp; circular; %GC: 36.5 5. CP185969.1 : plasmid; 41,501 bp; circular; %GC: 32.5 6. CP185970.1 : plasmid; 32,086 bp; circular; %GC: 35.5 7. CP185971.1 : plasmid; 3,166 bp; circular; %GC: 33	MiLongA pipeline 1.0.1 (https://gitlab.com/bfr_bioinformatics/milonga), Unicycler v0.4.8 (11, 12), Pilon v1.24 (10), platon v1.6 (18), Geneious Prime 2020.2.2
species based on ANI	ANI coverage: 92% ANI value: 99.9 best ANI reference strain: <i>B. mobilis</i> 0711P9-1 (type strain)	BakCharak pipeline v3.1.6 (13) utilizing FastANI v1.33 (19) and the typestrain database from BTypers v3.4.0 (20)
<i>panC</i> clade	clade 2 (legacy as well as adjusted clade)	BakCharak pipeline v3.1.6 (13) utilizing abricate v1.0.1 (https://github.com/tseemann/abricate) and the <i>panC</i> databases of (20)
signature of 16s rRNA gene and <i>cspA</i> gene	16s rRNA gene: mesophilic signature; <i>cspA</i> gene: mesophilic signature	BakCharak pipeline v3.1.6 (13) utilizing bakta v1.9.3 (21) and the mesophilic and psychrotolerant signature sequences according to (22) for <i>cspA</i> and (23) for the 16 s rRNA gene
MLST	784	BakCharak pipeline v3.1.6 (13) utilizing MLST v2.23.0 (https://github.com/tseemann/mlst) and the PubMLST Database: 2024-01-16 (24)
<i>cesHPTABCD</i> gene cluster	location: 349,073 bp plasmid coverage (%) and identity (%) with respective gene sequences from reference strain AH187: <i>cesH</i> : 100 and 100 <i>cesP</i> : 100 and 100 <i>cesT</i> : 100 and 100 <i>cesA</i> : 100 and 100 <i>cesB</i> : 100 and 100 <i>cesC</i> : 100 and 99.89 <i>cesD</i> : 100 and 100	BakCharak pipeline v3.1.6 (13) utilizing abricate v1.0.1 (https://github.com/tseemann/abricate) and the VFDB (SetB): 2022-08-26 (25)

github.com/rrwick/Porechop) and further processed and evaluated using NanoFilt v2.8.0 and NanoStat v1.5.0 (9). The short-read and the long-read data sets were *de novo* assembled and circularized using Unicycler v0.4.8, including Pilon for polishing (10–12).

Based on this assembly, we performed further characterizations with the pipeline BakCharak v3.1.6 (13). The annotation was completed with PGAP (14–16) at the NCBI.

Characterization results, along with details about read quality control, assembly, and applied tools, are summarized in Table 1.

AUTHOR AFFILIATION

¹Department Biological Safety, German Federal Institute for Risk Assessment (BfR), Berlin, Germany

AUTHOR ORCIDs

Hendrik Frentzel  <http://orcid.org/0009-0000-3551-0527>

AUTHOR CONTRIBUTIONS

Hendrik Frentzel, Data curation, Investigation, Writing – original draft | Ylanna Kelner-Burgos, Data curation, Investigation | Pascal Witt, Data curation, Investigation | Maria Borowiak, Data curation, Methodology, Software, Writing – review and editing | Carlus Deneke, Data curation, Methodology, Software, Writing – review and editing

DATA AVAILABILITY

All data are available under BioProject [PRJNA1037718](https://ncbi.nlm.nih.gov/bioproject/PRJNA1037718) and BioSample [SAMN38195819](https://ncbi.nlm.nih.gov/biosample/SAMN38195819). Sequencing raw reads were deposited in the NCBI Sequence Read Archive (SRA) under the accession numbers [SRX28127605](https://ncbi.nlm.nih.gov/sra/SRX28127605) (ONT data) and [SRX28127604](https://ncbi.nlm.nih.gov/sra/SRX28127604) (Illumina data). The complete genome sequence is available at NCBI GenBank under accession number [GCA_033841795.2](https://ncbi.nlm.nih.gov/genbank/GCA_033841795.2).

REFERENCES

- Carroll LM, Wiedmann M, Kovac J. 2020. Proposal of a taxonomic nomenclature for the *Bacillus cereus* group which reconciles genomic definitions of bacterial species with clinical and industrial phenotypes. *mBio* 11:e00034-20. <https://doi.org/10.1128/mBio.00034-20>
- Ehling-Schulz M, Frenzel E, Gohar M. 2015. Food-bacteria interplay: pathometabolism of emetic *Bacillus cereus*. *Front Microbiol* 6:704. <https://doi.org/10.3389/fmicb.2015.00704>
- Liu Y, Du J, Lai QL, Zeng RY, Ye DZ, Xu J, Shao ZZ. 2017. Proposal of nine novel species of the *Bacillus cereus* group. *Int J Syst Evol Microbiol* 67:2499–2508. <https://doi.org/10.1099/ijsem.0.001821>
- Hoton FM, Fornelos N, N'guessan E, Hu X, Swiecicka I, Dierick K, Jääskeläinen E, Salkinoja-Salonen M, Mahillon J. 2009. Family portrait of *Bacillus cereus* and *Bacillus weihenstephanensis* cereulide-producing strains. *Environ Microbiol Rep* 1:177–183. <https://doi.org/10.1111/j.1758-2229.2009.00028.x>
- Mei XF, Xu K, Yang LL, Yuan ZM, Mahillon J, Hu XM. 2014. The genetic diversity of cereulide biosynthesis gene cluster indicates a composite transposon Tncs in emetic *Bacillus weihenstephanensis*. *BMC Microbiol* 14:149. <https://doi.org/10.1186/1471-2180-14-149>
- Thorsen L, Hansen BM, Nielsen KF, Hendriksen NB, Phipps RK, Budde BB. 2006. Characterization of emetic *Bacillus weihenstephanensis*, a new cereulide-producing bacterium. *Appl Environ Microbiol* 72:5118–5121. <https://doi.org/10.1128/AEM.00170-06>
- Frentzel H, Kraemer M, Kelner-Burgos Y, Uelze L, Bodi D. 2024. Cereulide production capacities and genetic properties of 31 emetic *Bacillus cereus* group strains. *Int J Food Microbiol* 417:110694. <https://doi.org/10.1016/j.ijfoodmicro.2024.110694>
- Deneke C, Brendebach H, Uelze L, Borowiak M, Malorny B, Tausch SH. 2021. Species-specific quality control, assembly and contamination detection in microbial isolate sequences with AQUAMIS. *Genes (Basel)* 12:644. <https://doi.org/10.3390/genes12050644>
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl 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>
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom* 3:e000132. <https://doi.org/10.1099/mgen.0.000132>
- Deneke C, Strohlein A, Brendebach H. 2019. BakCharak - bacterial WGS characterization pipeline. Zenodo. <https://doi.org/10.5281/zenodo.15487524>
- Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, et al. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, et al. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>
- Chen SF, Zhou YQ, Chen YR, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Schwengers O, Barth P, Falgenhauer L, Hain T, Chakraborty T, Goesmann A. 2020. Platon: identification and characterization of bacterial plasmid contigs in short-read draft assemblies exploiting protein sequence-based replicon distribution scores. *Microb Genom* 6:mgen000398. <https://doi.org/10.1099/mgen.0.000398>
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>
- Carroll LM, Cheng RA, Kovac J. 2020. No assembly required: using BType3 to assess the congruency of a proposed taxonomic framework for the *Bacillus cereus* group with historical typing methods. *Front Microbiol* 11:580691. <https://doi.org/10.3389/fmicb.2020.580691>
- Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* 7:000685. <https://doi.org/10.1099/mgen.0.000685>
- Francis KP, Mayr R, von Stetten F, Stewart GS, Scherer S. 1998. Discrimination of psychrotrophic and mesophilic strains of the *Bacillus cereus* group by PCR targeting of major cold shock protein genes. *Appl Environ Microbiol* 64:3525–3529. <https://doi.org/10.1128/AEM.64.9.3525-3529.1998>
- von Stetten F, Francis KP, Lechner S, Neuhaus K, Scherer S. 1998. Rapid discrimination of psychrotolerant and mesophilic strains of the *Bacillus cereus* group by PCR targeting of 16S rDNA. *J Microbiol Methods* 34:99–106. [https://doi.org/10.1016/S0167-7012\(98\)00077-3](https://doi.org/10.1016/S0167-7012(98)00077-3)
- Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 3:124. <https://doi.org/10.12688/wellcomeopenres.14826.1>
- Liu B, Zheng D, Jin Q, Chen L, Yang J. 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res* 47:D687–D692. <https://doi.org/10.1093/nar/gky1080>