



# Genome Sequence of *Geobacillus stearothermophilus* DSM 458, an Antimicrobial-Producing Thermophilic Bacterium, Isolated from a Sugar Beet Factory

Kevin Egan,<sup>a</sup> Philip Kelleher,<sup>a</sup> Des Field,<sup>a,b</sup> Mary C. Rea,<sup>b,c</sup> R. Paul Ross,<sup>a,b</sup> Paul D. Cotter,<sup>b,c</sup> Colin Hill<sup>a,b</sup>

School of Microbiology, University College Cork, Cork City, Ireland<sup>a</sup>; APC Microbiome Institute, University College Cork, Cork City, Ireland<sup>b</sup>; Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland<sup>c</sup>

**ABSTRACT** This paper reports the full genome sequence of the antimicrobial-producing bacterium *Geobacillus stearothermophilus* DSM 458, isolated in a sugar beet factory in Austria. *In silico* analysis reveals the presence of a number of novel bacteriocin biosynthetic genes.

*Geobacillus* spp. are thermophilic, Gram-positive, aerobic or facultative aerobic, spore-forming bacteria, and are highly resistant to heat when in spore form (1). The genus *Geobacillus* and its species have been shown in recent years to produce a number of bacteriocins or bacteriocin-like inhibitory substances (BLIS) (2–4). Here, we report the full genome sequence of *Geobacillus stearothermophilus* DSM 458, which was isolated from sugar beet juice obtained from extraction installations in Austria. It has been shown to produce a narrow-spectrum antimicrobial substance that targets other geobacilli, the basis of which requires further characterization (5). In order to identify the gene(s) responsible for the production of this antimicrobial substance, we completed full-genome sequencing of this bacterium, for which the 16S rRNA sequence had previously been determined (GenBank accession number AY608931).

Cells of *G. stearothermophilus* DSM 458 were grown to mid-log phase in brain heart infusion (BHI) broth and centrifuged at 5,000 rpm for 20 min. A 600-mg pellet of cells was then snap frozen by placing the centrifuge tube into ethanol which had been previously cooled to  $-80^{\circ}\text{C}$ . Chromosomal DNA was isolated by commercial sequence provider GATC Biotech, Ltd. (Konstanz, Germany). Single-molecule real-time (SMRT) sequencing was performed on a Pacific Biosciences RS II sequencing platform (executed by GATC Biotech, Ltd., Germany) to a mean fold coverage of 147.88. *De novo* assembly of the genome was performed using the SMRTPortal analysis platform (version 2.3.1), utilizing the RS\_HGAP\_Assembly.2 protocol. This resulted in a single contiguous chromosome of 3,466,824 bp and a G+C content of 52.11%.

Following assembly of the genome, open reading frame (ORF) prediction was performed using the Prodigal v2.5 prediction software (6) and confirmed using BLASTx alignments (7). The genome was then automatically annotated using BLASTp (7) against the nonredundant protein database curated by the National Center for Biotechnology Information (NCBI). In addition to this, manual curation of the genome was verified using the Artemis genome browser and annotation tool (8), i.e., inspection of ORF results, adjustment of start codons where necessary, and identification of pseudogenes. Further genomic analysis utilized the programs tRNAscan-SE (9), RNAmmer (10), PHAST (11), and CRISPRFinder (12), showing that the genome contains 3,525 genes, 3,361 proteins, 43 pseudogenes, 32 rRNAs, 89 tRNAs, 4 putative phages, and 5

Received 18 September 2017 Accepted 18 September 2017 Published 26 October 2017

**Citation** Egan K, Kelleher P, Field D, Rea MC, Ross RP, Cotter PD, Hill C. 2017. Genome sequence of *Geobacillus stearothermophilus* DSM 458, an antimicrobial-producing thermophilic bacterium, isolated from a sugar beet factory. *Genome Announc* 5:e01172-17. <https://doi.org/10.1128/genomeA.01172-17>.

**Copyright** © 2017 Egan 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 Colin Hill, [c.hill@ucc.ie](mailto:c.hill@ucc.ie).

clustered regularly interspaced short palindromic repeat (CRISPR)-associated repeat regions. Additionally, BAGEL3 (13) predicts the presence of a novel circularin-like bacteriocin with a 46% identity to circularin A (14), as determined by pairwise alignment using BLASTp (7). AntiSMASH v4.0.2 (15) also highlighted the presence of this circularin-like bacteriocin and a gene cluster encoding a bacteriocin with homology to Linocin M18 (16). The Linocin M18-like bacteriocin was determined to have a 25% identity to the Linocin M18 protein sequence, as determined by pairwise alignment using BLASTp (7).

Combining the *in silico* analysis with *in vitro* experimentation should ultimately lead to characterization and greater understanding of the antimicrobial potential/nature of this strain. Furthermore, the availability of this genome will ultimately contribute to the accuracy of *in silico* prediction software for the genus of *Geobacillus* in the future.

**Accession number(s).** The full genome sequence of *Geobacillus stearothermophilus* DSM 458 has been deposited in GenBank under the accession number [CP016552](https://www.ncbi.nlm.nih.gov/nuccore/CP016552).

## ACKNOWLEDGMENTS

K.E., D.F., M.C.R., R.P.R., P.D.C., and C.H. were supported and funded by the Irish Government under the National Development Plan, through the Food Institutional Research Measure (DAFM 13/F/462), administered by the Department of Agriculture, Fisheries and Food, Ireland. P.K., D.F., M.C.R., R.P.R., P.D.C., and C.H. were supported by the APC Microbiome institute under grant SFI/12/RC/2273. P.D.C. was also funded by a SFI PI award under grant 11/PI/1137. R.P.R. and C.H. were also supported by SFI investigator awards under grant 10/IN.1/B3027.

We also thank Daniel Zeigler for the kind donation of this strain from the BGSC (*Bacillus* Genetic Stock Center).

## REFERENCES

- Nazina TN, Tourova TP, Poltarau AB, Novikova EV, Grigoryan AA, Ivanova AE, Lysenko AM, Petrunyaka VV, Osipov GA, Belyaev SS, Ivanov MV. 2001. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenuatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus*, *G. th.* *Int J Syst Evol Microbiol* 51:433–446. <https://doi.org/10.1099/00207713-51-2-433>.
- Garg N, Tang W, Goto Y, Nair SK, van der Donk WA. 2012. Lantibiotics from *Geobacillus thermodenitrificans*. *Proc Natl Acad Sci U S A* 109: 5241–5246. <https://doi.org/10.1073/pnas.1116815109>.
- Özdemir GB, Biyik HH. 2012. Isolation and characterization of toebicin 218, a bacteriocin, produced by *Geobacillus toebii* HBB-218. *Afr J Biotechnol* 11:7711. <https://doi.org/10.5897/AJB11.1331>.
- Pokusaeva K, Kuisiene N, Jasinskyte D, Rutiene K, Saleikiene J, Chitavichius D. 2009. Novel bacteriocins produced by *Geobacillus stearothermophilus*. *Cent Eur J Biol* 4:196–203. <https://doi.org/10.2478/s11535-009-0009-1>.
- Egan K, O'Connor PM, Field D, Hill C, Cotter PD. 2016. Poster abstract A11. *Geobacillus stearothermophilus*: A potential source of bacteriocins? AMP2016 International Symposium on Antimicrobial Peptides, 6 to 8 June 2016, Montpellier, France.
- Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
- 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](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <https://doi.org/10.1093/bioinformatics/16.10.944>.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964.
- Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <https://doi.org/10.1093/nar/gkm160>.
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. <https://doi.org/10.1093/nar/gkr485>.
- Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52–W57. <https://doi.org/10.1093/nar/gkm360>.
- van Heel AJ, de Jong A, Montalbán-López M, Kok J, Kuipers OP. 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Res* 41:W448–W453. <https://doi.org/10.1093/nar/gkt391>.
- Kawai Y, Kemperman R, Kok J, Saito T. 2004. The circular bacteriocins gassericin A and circularin A. *Curr Protein Pept Sci* 5:393–398. <https://doi.org/10.2174/1389203043379549>.
- Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* 43:W237–W243. <https://doi.org/10.1093/nar/gkv437>.
- Valdés-Stauber N, Scherer S. 1994. Isolation and characterization of Linocin M18, a bacteriocin produced by *Brevibacterium linens*. *Appl Environ Microbiol* 60:3809–3814.