Draft Genome Sequence of *Bacillus velezensis* CE2, Which Genetically Encodes a Novel Multicomponent Lantibiotic

Emily Campbell,[a] Michelle Gerst,[b] B. Carol Huang,[c] Nguyet Kong,[c] Bart C. Weimer,[c] Ahmed E. Yousef[a,b]

[a]Department of Food Science and Technology, The Ohio State University, Columbus, Ohio, USA
[b]Department of Microbiology, The Ohio State University, Columbus, Ohio, USA
[c]School of Veterinary Medicine, 100K Pathogen Genome Project, University of California Davis, Davis, California, USA

ABSTRACT *Bacillus velezensis* CE2 produces potent antimicrobial compound(s). The draft genome sequence of the strain reported here is 4.1 Mb with a G+C content of 46.1%. Whole-genome sequencing revealed that the strain genetically encodes a novel multicomponent lantibiotic, velezensicidin.

A number of *Bacillus velezensis* strains were reported to produce antimicrobials (1–6). A high-throughput culture-based assay was used to screen soil microbiota for antimicrobial production (7). Using this method, *B. velezensis* CE2 was isolated and found to produce potent antimicrobial activity against Gram-positive bacteria.

The whole genome of *B. velezensis* CE2 was sequenced within the 100K Pathogen Genome Sequencing Project (8), as described previously (9). Briefly, genomic DNA (gDNA) was extracted and purified from *B. velezensis* CE2 culture, grown on MRS agar (Becton, Dickinson, Franklin Lakes, NJ) at 30°C for 24 h, using a Qiagen DNA genomic minikit (catalog number 51306). gDNA quality was determined, using $A_{260/280}$ and $A_{260/230}$ ratios, to be >1.5, and size characterization was done using an Agilent 2200 TapeStation system, as described previously (10). Libraries were constructed using a Kapa HyperPlus library preparation kit (catalog number KK8514) and indexed with Weimer 384 TS-LT DNA barcodes (Integrated DNA Technologies, Inc., Coralville, IA, USA). The resulting libraries were 350 to 450 bp. Library amplification was completed in eight cycles using the Kapa HiFi HotStart ReadyMix PCR kit, followed by cleanup with 1× solid-phase reversible immobilization (SPRI) magnetic beads (9). Final library quality control was done using an Agilent 2100 Bioanalyzer with a high-sensitivity DNA kit. Sequencing library concentration was determined by quantitative PCR (qPCR) using the Kapa library quantification kit with universal qPCR mix (catalog number KK4824) before indexing. Libraries were sequenced using an Illumina HiSeq X Ten instrument with a 150-bp paired end (PE150) read (Novagene, Sacramento, CA). For *B. velezensis* CE2, the total number of paired reads generated was 1,213,887. These were assembled using SPAdes version 3.11.1 (settings: k-mer sizes 21, 33, 55, and 77 with mismatch careful mode) (11), after PhiX and indexes were removed by NCBI's contamination filter (VecScreen). After assembly, contigs were ordered using *Bacillus velezensis* FZB42 as a reference, with progressiveMAUVE (version 2.4.0) and applying the multiple genome alignment option (12). The genome contained 265 contigs, with a coverage of 44×. The maximum and minimum sizes of contigs were 558,902 bp and 225 bp, respectively, with an N_{50} value of 297,815. The genome was annotated by Rapid Annotations using Subsystems Technology (RAST) version 2.0 (default settings), and 4,438 coding regions and 88 rRNA were found (13).

BAGEL version 3 (default settings) was used to identify the lantibiotic gene cluster in the draft genome sequence (14). Two-component lantibiotic structural genes (*lanA*), two modification enzymes genes (*lanM*), and a transporter gene (*lanT*) were identified

in the genome. To the best of our knowledge, the amino acid sequences encoded by the structural genes for the two-component lantibiotic were not reported elsewhere in published literature. The production of potent antimicrobial activity and presence of a novel lantibiotic structural gene indicate the potential usefulness of *B. velezensis* CE2 in food or agriculture.

ACKNOWLEDGMENTS

REFERENCES

1. Palazzini JM, Dunlap CA, Bowman MJ, Chulze SN. 2016. *Bacillus velezensis* RC 218 as a biocontrol agent to reduce *Fusarium* head blight and deoxynivalenol accumulation: genome sequencing and secondary metabolite cluster profiles. Microbiol Res 192:30–36. https://doi.org/10.1016/j.micres.2016.06.002.
2. Lee H-H, Park J, Lim JY, Kim H, Choi GJ, Kim J-C, Seo Y-S. 2015. Complete genome sequence of *Bacillus velezensis* G341, a strain with a broad inhibitory spectrum against plant pathogens. J Biotechnol 211:97–98. https://doi.org/10.1016/j.jbiotec.2015.07.005.
3. Chen L. 2017. Complete genome sequence of *Bacillus velezensis* LM2303, a biocontrol strain isolated from the dung of wild yak inhabited Qinghai-Tibet plateau. J Biotechnol 251:124–127. https://doi.org/10.1016/j.jbiotec.2017.04.034.
4. Kim SY, Lee SY, Weon H-Y, Sang MK, Song J. 2017. Complete genome sequence of *Bacillus velezensis* M75, a biocontrol agent against fungal plant pathogens, isolated from cotton waste. J Biotechnol 241:112–115. https://doi.org/10.1016/j.jbiotec.2016.11.023.
5. Jin Q, Jiang Q, Zhao L, Su C, Li S, Si F, Li S, Zhou C, Mu Y, Xiao M. 2017. Complete genome sequence of *Bacillus velezensis* S3-1, a potential biological pesticide with plant pathogen inhibiting and plant promoting capabilities. J Biotechnol 259:199–203. https://doi.org/10.1016/j.jbiotec.2017.07.011.
6. Wang B, Liu H, Ma H, Wang C, Liu K, Li Y, Hou Q, Ge R, Zhang T, Liu F, Ma J, Wang Y, Wang H, Xu B, Yao G, Xu W, Fan L, Ding Y, Du B. 2017. Complete genome sequence of biocontroller *Bacillus velezensis* strain JTYP2, isolated from leaves of *Echeveria laui*. Genome Announc 5:e00505-17. https://doi.org/10.1128/genomeA.00505-17.
7. Gerst MM, Yousef AE. 2018. Modified microassay for the isolation of antimicrobial-producing, spore-forming and nonspore-forming bacteria. J Appl Microbiol 124:1401–1410. https://doi.org/10.1111/jam.13724.
8. Weimer BC. 2017. 100K Pathogen Genome Project. Genome Announc 5:e00594-17. https://doi.org/10.1128/genomeA.00594-17.
9. Kong N, Davis M, Arabyan N, Huang BC, Weis AM, Chen P, Thao K, Ng W, Chin N, Foutouhi S, Foutouhi A, Kaufman J, Xie Y, Storey DB, Weimer BC. 2017. Draft genome sequences of 1,183 Salmonella strains from the 100K pathogen genome project. Genome Announc 5:e00518-17. https://doi.org/10.1128/genomeA.00518-17.
10. Kong N, Ng W, Cai L, Leonardo A, Kelly L, Weimer BC. 2014. Integrating the DNA integrity number (DIN) to assess genomic DNA (gDNA) quality control using the Agilent 2200 TapeStation system. Application note. Agilent Technologies, Santa Clara, CA. http://www.agilent.com/cs/library/applications/5991-5442EN.pdf.
11. 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 A J Comput Mol Cell Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
12. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics 25:2071–2073. https://doi.org/10.1093/bioinformatics/btp356.
13. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
14. 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.