



Draft Genome Sequence of *Bacillus lehensis* M136, Isolated from the Manleluag Hyperalkaline Spring in Pangasinan, Philippines

Aprill P. Manalang,^a Andrew D. Montecillo,^a Lorren Myriccar G. Berdos,^a Rose Ann G. Franco,^a Eula Francia M. Bosito,^a Noel G. Sabino,^a Maria Genaleen Q. Diaz,^a Nacita B. Lantican^a

^aInstitute of Biological Sciences, University of the Philippines Los Baños College, Laguna, Philippines

ABSTRACT Here, we report the draft whole-genome sequence of *Bacillus lehensis* M136, isolated from a hyperalkaline spring located in Pangasinan, Philippines. From 24 scaffolds, the total genome assembly length is 3,985,437 bp. Industrially important genes like cyclodextrin glycosyltransferase (CGTase) and proteases were detected in this draft genome.

Bacillus lehensis, an alkalitolerant bacterium, was initially reported as a novel species by Ghosh et al. (1) following its isolation from soil samples in Leh, India. Its ability to produce enzymes such as cyclodextrin glycosyltransferase (CGTase) (2), maltogenic amylase (3), and protease (1) is of interest due to its many industrial applications. In this paper, we report the whole-genome sequence of *B. lehensis* M136 (C1-1), isolated from Manleluag Hyperalkaline Spring in Pangasinan, Philippines. The recorded pH range of the sampling site was pH 10 to 11.3, and the temperature ranged from 33 to 36°C. This strain is characterized by a “rotten egg” odor and blue flame ignition from gas bubble emissions, indicative of the high methane and sulfur content. The presence of serpentinizing seeps was also observed at the site. A water sample from the spring was collected and enriched using Horikoshi I medium, which generally supports growth of alkaliphilic and alkalitolerant bacteria. The putative *B. lehensis* M136 isolate, grown in Horikoshi I broth for 48 hours at 37°C, was sent for processing at the Philippine Genome Center DNA Sequencing Core Facility (Quezon City, Philippines). Genomic DNA was extracted using the Purelink genomic DNA extraction kit (Life Technologies Corporation, Carlsbad, CA, USA). Using 56.5 ng of DNA, a paired-end library was prepared using the Nextera XT library preparation kit and was sequenced using the Illumina MiSeq v3 2 × 300-cycle kit (Illumina, Inc., San Diego, CA, USA).

A total of 733,980 raw paired-end reads were generated. Quality trimming, carried out with the Trimmomatic tool (4) (using the following parameters: leading, 3; trailing, 3; slidingwindow, 4:15; minlen, 36; and headcrop, 10), produced 656,232 paired-end reads with a length of 26 to 291 bp. *De novo* assembly was performed using SPAdes version 3.11.1 (5) following default parameters and the “–careful” option to reduce the number of mismatches and short indels. Initial assembly produced 131 contigs, with the largest contig length of 1,848,692 bp and an N_{50} of 846,478 bp. To scaffold and gap fill the assembly, assembly improvement software (6) was used with default parameters. The assembly quality was analyzed using QUAST 4.6 (7), after which contigs with less than 500 bp and low coverage ($<2.0\times$) were removed. The taxonomic classification of the isolate was established through Microbial Genome Atlas (MiGA) (8) analysis using the NCBI RefSeq and Prokaryotic databases. The contigs were then reordered using Mauve (9) Contig Mover with *B. lehensis* G1 NZ (GenBank accession number CP003923) as the reference. The draft genome sequence was annotated using the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP)

Citation Manalang AP, Montecillo AD, Berdos LMG, Franco RAG, Bosito EFM, Sabino NG, Diaz MGQ, Lantican NB. 2019. Draft genome sequence of *Bacillus lehensis* M136, isolated from the Manleluag Hyperalkaline Spring in Pangasinan, Philippines. Microbiol Resour Announc 8:e00029-19. <https://doi.org/10.1128/MRA.00029-19>.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Copyright © 2019 Manalang 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 Aprill P. Manalang, apmanalang@up.edu.ph.

Received 17 January 2019

Accepted 27 February 2019

Published 28 March 2019

(10) and the Rapid Annotations using Subsystems Technology (RAST) version 2 RASTtk pipeline (11, 12).

The *B. lehensis* M136 genome was found to contain 24 scaffolds and 29 contigs, including 5 gaps, for a total of 3,985,437 bp and a 39.8% G+C content (N_{50} , 846,478 bp; coverage, 216×). Based on MiGA, the genome was found to be most similar to the genome of *Bacillus lehensis* G1 NZ (CP003923) (99.12% average nucleotide identity [ANI]), which was used in the contig reordering.

The NCBI PGAP predicted 3,923 protein-coding sequences, 14 rRNAs, 72 tRNAs, and 5 noncoding RNAs (ncRNAs). RASTtk annotation revealed the presence of genes for industrially important enzymes like CGTase (EC 2.4.1.19) and several proteases in this draft genome of *B. lehensis* M136.

Data availability. The genome sequence of *Bacillus lehensis* M136 has been deposited at DDBJ/ENA/GenBank under the accession number [RQRY00000000](https://www.ncbi.nlm.nih.gov/nuclseq/RQRY00000000) and SRA accession number [PRJNA506874](https://www.ncbi.nlm.nih.gov/sra/PRJNA506874). The version described in this paper is version RQRY01000000.

ACKNOWLEDGMENTS

This research was supported by the University of the Philippines System Emerging Interdisciplinary Research (EIDR) Program and the Philippine Council for Industry, Energy and Emerging Technology Research and Development (PCIEERD) of the Department of Science and Technology (DOST) as part of the research program titled: "Molecular and Culture-based Characterization of Bacterial Community from Manleluag Hyperalkaline Spring in Pangasinan."

We also thank M. Anita Bautista of the Philippine Genome Center (PGC) for facilitating the sequencing of our samples.

REFERENCES

1. Ghosh A, Bhardwaj M, Satyanarayana T, Khurana M, Mayilraj S, Jain RK. 2007. *Bacillus lehensis* sp. nov., an alkalitolerant bacterium isolated from soil. *Int J Syst Evol Microbiol* 57:238–242. <https://doi.org/10.1099/ijs.0.64617-0>.
2. Abdul Manas NH, Pachelles S, Mahadi NM, Illias RM. 2014. The characterization of an alkali-stable maltogenic amylase from *Bacillus lehensis* G1 and improved malto-oligosaccharide production by hydrolysis suppression. *PLoS One* 9:e106481. <https://doi.org/10.1371/journal.pone.0106481>.
3. Blanco KC, de Lima CJB, Monti R, Martins J, Jr, Bernardi NS, Contiero J. 2012. *Bacillus lehensis*—an alkali-tolerant bacterium isolated from cassava starch wastewater: optimization of parameters for cyclodextrin glycosyltransferase production. *Ann Microbiol* 62:329–337. <https://doi.org/10.1007/s13213-011-0266-x>.
4. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
5. 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>.
6. Page AJ, De Silva N, Hunt M, Quail MA, Parkhill J, Harris SR, Otto TD, Keane JA. 2016. Robust high-throughput prokaryote *de novo* assembly and improvement pipeline for Illumina data. *Microbial Genomics* 2:e000083. <https://doi.org/10.1099/mgen.0.000083>.
7. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
8. Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT. 2018. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Res* 46:W282–W288. <https://doi.org/10.1093/nar/gky467>.
9. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
10. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufu S, Li W. 2013. Prokaryotic genome annotation pipeline. In Beck J, Benson D, Coleman J, Hoepfner M, Johnson M, Maglott D, Mizrahi I, Morris R, Ostell J, Pruitt K, Rubenstein W, Sayers E, Sirotkin K, Tatusova T (ed), *The NCBI handbook*, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.
11. 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>.
12. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>.