



Draft Genome Sequence of *Bacillus thuringiensis* Strain UNMSM10RA, Isolated from Potato Crop Soil in Peru

Yolanda Bedsabé Delgado-Silva,^a David Tarazona,^a Fernando Serna,^{a,b} Eduardo Juscamayta,^c Julio César Chávez-Galarza,^{a,d} Evelyn Roxana Farfán-Vignolo,^a Gabriel Delgado,^a Abad Flores,^c Gabriela Solano,^c Dina L. Gutierrez^a

^aDirección de Recursos Genéticos y Biotecnología, Instituto Nacional de Innovación Agraria (INIA), La Molina, Lima, Perú

^bDirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), La Molina, Lima, Perú

^cLaboratorio de Microbiología Ambiental y Biotecnología, Facultad de Ciencias Biológicas, Universidad Nacional Mayor de San Marcos, Ciudad Universitaria, Lima, Perú

^dEscuela Académico Profesional de Agronomía, Universidad Nacional de Cañete, Urb. Santa Rosa de Hualcará, San Vicente de Cañete, Lima, Perú

ABSTRACT The 5.5-Mb genome sequence of *Bacillus thuringiensis* strain UNMSM10RA, isolated from potato crop soil, is reported in this study. The strain UNMSM10RA contains 5,347 protein-coding sequences, 105 tRNA genes, 15 rRNA genes, and 5 noncoding RNA (ncRNA) genes, with an average G+C content of 35.1%.

Bacillus thuringiensis is an aerobic, ubiquitous, rod-shaped, Gram-positive, spore-forming bacterium that has been isolated from a variety of environments. This bacterium is well known for producing parasporal crystalline inclusions containing Cry and Cyt proteins (also known as δ -endotoxins) (1). Currently, the sequencing of *Bacillus thuringiensis* strain genomes has increased and facilitated the search of not only new molecules with insecticidal activity but also genes conferring the ability for heavy metal biosorption, resistance, or tolerance (2).

We isolated *Bacillus thuringiensis* strains from potato crop soils in Huaral, Lima, Peru. Serial dilutions (10^{-3}) were done for each sample in a peptone salt solution and incubated at 70°C for 15 min, followed by thermal shock. Then, the samples were incubated at 70°C for 20 min, followed by incubation at 30°C for 18 h. Positives samples were cultured overnight at 30°C in Luria-Bertani (LB) medium under aerobic conditions. One of the colonies recovered, named UNMSM10RA, was chosen for sequencing.

For UNMSM10RA genome sequencing, total genomic DNA was extracted from overnight cultures using the High Pure PCR template preparation kit (Roche, USA) according to the manufacturer's instructions. Genomic DNA was used for library preparation with the Nextera XT DNA library preparation kit. The library showed a size range of ~250 to 1,000 bp. Whole-genome sequencing was performed on an Illumina MiSeq system (San Diego, CA) in a paired-end mode with a read length of 150 bp. A total of 969,034 reads were generated. Quality control of these reads was performed with the FastQC tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). *De novo* genome assembly was done using Velvet v.1.2.10 (3) and SPAdes v.3.12.1 (4) and merged for the best assembly with genomic assemblies merger for next-generation sequencing (GAM-NGS) through the PATRIC server on 22 January 2018 (5). Genome annotation was carried out using the Rapid Annotations using Subsystems Technology (RAST) platform (6), NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP), and Clusters of Orthologous Groups (7). tRNA and rRNA were predicted by tRNAscan-SE (8) and RNAmmer (9), respectively. *Bacillus* species confirmation was performed by the comparison of genome-derived 16S rRNA sequences and establishing their phylogenetic relationships using MEGA X software (10). All programs were set up with default parameters.

A comparison of 16S rRNA sequences revealed 100% similarity between *Bacillus*

Citation Delgado-Silva YB, Tarazona D, Serna F, Juscamayta E, Chávez-Galarza JC, Farfán-Vignolo ER, Delgado G, Flores A, Solano G, Gutierrez DL. 2020. Draft genome sequence of *Bacillus thuringiensis* strain UNMSM10RA, isolated from potato crop soil in Peru. *Microbiol Resour Announc* 9:e01189-19. <https://doi.org/10.1128/MRA.01189-19>.

Editor Irene L. G. Newton, Indiana University, Bloomington

Copyright © 2020 Delgado-Silva 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 Yolanda Bedsabé Delgado-Silva, ybdelgados@gmail.com, or Dina L. Gutierrez, dgutierrez@inia.gob.pe.

Received 14 October 2019

Accepted 4 December 2019

Published 9 January 2020

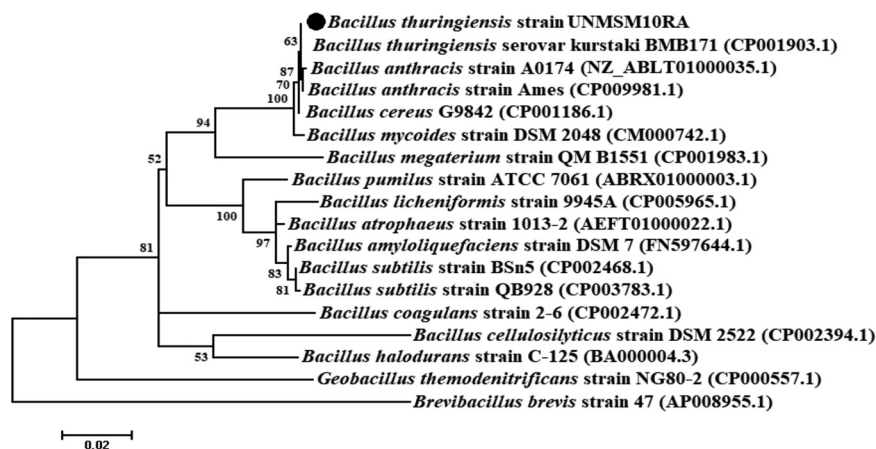


FIG 1 Phylogenetic relationships based on 16S rRNA sequences highlighting the position of *Bacillus thuringiensis* strain UNMSM10RA and other species of *Bacillus* and related genera. GenBank accession numbers are presented in parentheses. This topology was inferred from a multiple alignment using ClustalW 2.1 and the maximum likelihood method within MEGA X. Nodes are supported by 500 bootstraps.

thuringiensis strain UNMSM10RA and *Bacillus thuringiensis* serovar kurstaki BMB171, and this was corroborated by phylogenetic proximity (Fig. 1). The estimated genome size of *B. thuringiensis* strain UNMSM10RA was 5,524,132 bp distributed in 449 contigs, with an N_{50} value of 61,078 bp and an average G+C content of 35.1%. Genome analysis revealed 5,347 coding sequences, 105 tRNA genes, 15 rRNA genes, 5 noncoding RNA (ncRNA) genes, and 713 pseudogenes. Within protein-coding genes, 31 were detected and identified in the metabolism of heavy metals such as arsenic, cadmium, and copper. Further analyses will be performed and provide more information for understanding the evolution of these *Bacillus thuringiensis* strains and their potential uses.

Data availability. This whole-genome shotgun project has been deposited in NCBI/GenBank under BioProject number [PRJNA448338](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA448338), BioSample number [SAMN08828790](https://www.ncbi.nlm.nih.gov/biosample/SAMN08828790), GenBank accession number [NZ_PZQR00000000](https://www.ncbi.nlm.nih.gov/genbank/NZ_PZQR00000000), and SRA number [SRR10158671](https://www.ncbi.nlm.nih.gov/sra/SRR10158671). The version described in this paper is the first version, NZ_PZQR01000000.

REFERENCES

- Palma L, Muñoz D, Berry C, Murillo J, Caballero P. 2014. *Bacillus thuringiensis* toxins: an overview of their biocidal activity. *Toxins* (Basel) 6:3296–3325. <https://doi.org/10.3390/toxins6123296>.
- Yik LY, Chin G, Budiman C, Joseph CG, Musta B, Rodrigues KF. 2018. Adaptive strategies of *Bacillus thuringiensis* isolated from acid mine drainage site in Sabah, Malaysia. *Indian J Microbiol* 58:165–173. <https://doi.org/10.1007/s12088-017-0701-1>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
- 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>.
- Wattam AR, Brettin T, Davis JJ, Gerdes S, Kenyon R, Machi D, Mao C, Olson R, Overbeek R, Pusch GD, Shukla MP, Stevens R, Vonstein V, Warren A, Xia F, Yoo H. 2018. Assembly, annotation, and comparative genomics in PATRIC, the all bacterial Bioinformatics Resource Center. *Methods Mol Biol* 1704:79–101. https://doi.org/10.1007/978-1-4939-7463-4_4.
- 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>.
- Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28:33–36. <https://doi.org/10.1093/nar/28.1.33>.
- 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. <https://doi.org/10.1093/nar/25.5.955>.
- 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>.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>.