Draft Genome Sequence of *Bacillus toyonensis* Strain GM18, Isolated from Agricultural Soil

Carlos J. Moreno-Peña,[a] Jorge H. García-García,[a] Cesar I. Hernández-Vásquez,[a] Luis J. Galán-Wong,[a] Benito Pereyra-Alférez[a]

aInstituto de Biotecnología, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, San Nicolas de los Garza, Nuevo León, Mexico

ABSTRACT *Bacillus toyonensis* is a recently described species related to *Bacillus cereus* and *Bacillus thuringiensis*. The GM18 strain previously identified as *B. thuringiensis* is now classified as *B. toyonensis* based on the RNA 16S sequence and whole-genome average nucleotide identity. The genome analysis revealed the presence of insecticide, nematicide, and antitumoral proteins.

The GM18 bacterium was isolated from agricultural soil collected in the state of Guanajuato, Mexico, using the methodology described by Travers et al. (1), and based on its H antigen characteristics, this strain was recognized as a new serotype (H-24) and named *Bacillus thuringiensis* subsp. *neolonensis*. In an early description, Rodriguez-Padilla et al. (2) showed that GM18 possesses characteristics of the *B. thuringiensis* group; it is Gram positive and sporulated and synthesizes a parasporal crystal. However, it had no insecticidal activity. Later, when the inclusion body was tested against cancer cell lines, solubilized proteins displayed biological activity, and cytotoxicity ranged from low against 1c1c7 and A431 to high against A375 and BpRc1 cell lines, but no cytotoxicity for sheep erythrocytes was detected (3).

For the genome sequencing of *B. toyonensis* GM18, cells obtained from our laboratory cryostock were grown in 25 ml of LB broth for 6 h at 30°C, and then the pellet was recovered after centrifugation and washed with sterile water. The pellet was shipped chilled to Genewiz (New Jersey, USA), where total DNA extraction was performed, and the DNA library was prepared with the Illumina DNA prep, (M) tagmentation, and IDT for Illumina DNA/RNA indexes set A tagmentation kits (Illumina, Inc., California, USA). Whole-genome sequencing of *B. toyonensis* GM18 was performed with paired-end Illumina MiSeq sequencing, which resulted in 11,175,537 paired-end reads of 137 to 150 bp with a coverage of 250×; for genome assembly and annotation, we used Galaxy (4). In brief, the sequence quality was determined with FastQC (Galaxy version 0.72+galaxy1) (5). Then, low-quality sequences and adapters were removed with Trimmomatic (Galaxy version 0.038.0) (6). The processed sequence reads were assembled with the Shovill pipeline, including the SPAdes assembler (Galaxy version 1.1.0+galaxy0) (7); the resulting contigs were annotated with Prokka (Galaxy version 1.14.5+galaxy0) (8) and with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) using the best-placed reference protein set method (GeneMarkS-2+) (9, 10); 91.59% of all sequences were assembled in 156 contigs with an N_{50} value of 171,900 bp. Default parameters were used for all software.

We found that GM18 has a 6.13-Mb genome containing 6,129 open reading frames (ORF) with a GC content of 34.93%. The genome has 96 operons for tRNAs and 16 for rRNAs. The average nucleotide identity (ANI) of the whole genome determined during the GenBank submission process (11) showed 99.018% identity with the type genome of *Bacillus toyonensis*; this and the sequence identity of the rRNA 16S (GenBank accession number JADNYS010000095.1), *gyrB* (MBF7148837.1), *groEL* (MBF7150603.1), and

XRE (MBF7150840.1) genes suggests that this isolate belongs to the *B. toyonensis* species (12).

In brief, the sequence of the GM18 genome could be key to understanding the *Bacillus* group and its members' evolutionary relations. This genome shows greater identity with *Bacillus toyonensis*, although it has protein-coding genes with identity to cry (GenBank accession number WP_195755641.1, WP_195755715.1), vip (WP_195755724.1, WP_195755747.1), and parasporin (WP_195755716.1) that are characteristic of *B. thuringiensis* species; it also has insertion sequences with identity to those of *B. thuringiensis*, but this could be because *B. toyonensis* is a relatively newly described species of the *Bacillus* group (13).

Data availability. The *Bacillus toyonensis* GM18 whole-genome shotgun sequence has been deposited in GenBank under the accession number JADNYS000000000 and BioProject number PRJNA676032. The raw sequence files can be found under SRA number SRP301246.

REFERENCES

1. Travers RS, Martin PA, Reichelderfer CF. 1987. Selective process for efficient isolation of soil *Bacillus* spp. Appl Environ Microbiol 53:1263–1266. https://doi.org/10.1128/AEM.53.6.1263-1266.1987.
2. Rodriguez-Padilla C, Galan-Wong L, de Barjac H, Roman-Calderon E, Tamez-Guerra R, Dulmage H. 1990. *Bacillus thuringiensis* subspecies neoleonensis serotype H-24, a new subspecies which produces a triangular crystal. J Invertebr Pathol 56:280–282. https://doi.org/10.1016/0022-2011(90)90111-l.
3. Espino-Vázquez AN. 2014. Caracterización biológica de parasporinas en cepas nativas de Bacillus thuringiensis. PhD thesis. Universidad Autónoma de Nuevo León, Nuevo León, Mexico. http://eprints.uanl.mx/id/eprint/4042.
4. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. 2010. Galaxy: a Web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol 89:19.10.1–19.10.21. https://doi.org/10.1002/0471142727.mb1910s89.
5. Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
6. 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.
7. Seemann, T. 2017. Shovill: faster SPAdes assembly of Illumina reads. https://github.com/tseemann/shovill.
8. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
9. 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.
10. 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, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068.
11. Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int J Syst Evol Microbiol 68:2386–2392. https://doi.org/10.1099/ijsem.0.002809.
12. Chelliah R, Wei S, Park BJ, Rubab M, Banan-Mwine Daliri E, Barathikannan K, Jin YG, Oh DH. 2019. Whole genome sequence of *Bacillus thuringiensis* ATCC 10792 and improved discrimination of *Bacillus thuringiensis* from *Bacillus cereus* group based on novel biomarkers. Microb Pathog 129:284–297. https://doi.org/10.1016/j.micpath.2019.02.014.
13. Jiménez G, Urdiain M, Cifuentes A, López-López A, Blanch AR, Tamames J, Kämpfer P, Kolstø AB, Ramón D, Martínez JF, Codoñer FM, Rosselló-Móra R. 2013. Description of *Bacillus toyonensis* sp. nov., a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the species of the group by means of ANI calculations. Syst Appl Microbiol 36:383–391. https://doi.org/10.1016/j.syapm.2013.04.008.