Draft Genome Sequence of Heavy Metal-Resistant *Bacillus cereus* NWUAB01

Olubukola Oluranti Babalola,[a] Bukola Rhoda Aremu,[a] Ayansina Segun Ayangbenro[a]

aFood Security and Safety Niche, Faculty of Natural and Agricultural Sciences, North-West University, Mmabatho, South Africa

ABSTRACT *Bacillus cereus* NWUAB01 was isolated from a gold-mining site in Vryburg, South Africa, for its multiple heavy metal resistance properties. Here, we report the draft genome sequence of *B. cereus* NWUAB01 obtained with Illumina sequencing.

Bacteria from polluted environments have evolved resistance and adaptive genes required for their survival in hostile environments. The search for novel molecular mechanisms from polluted environments provides strong informational resources that lead to new technologies (1). As a result, the draft genome of *Bacillus cereus* NWUAB01 isolated from soil of a gold-mining site in Vryburg, South Africa (26.1681354 S, 25.246582 E), is reported here.

Bacillus cereus is a spore-forming Gram-positive bacterium that is ubiquitous in nature and always detected in hostile environments (2). The spore-forming ability helps it to survive in different habitats (3). *B. cereus* NWUAB01 is a metal-resistant spore-forming isolate. Microbial degradation is a major route of recovery of pollutants from polluted environments by transformation and detoxification mechanisms (4). *B. cereus* NWUAB01 was identified from 98 bacterial isolates from soil samples obtained from a gold-mining site. The collection was screened for heavy metal (Cd, Cr, and Pb) resistance and secondary metabolite production. Soil samples were serially diluted and plated out on Luria-Bertani (LB) agar that had been supplemented with 50 mg/liter of each metal solution and a mixture of all the metals. The metal solutions ($CdSO_4$, K_2CrO_4, and $Pb(NO_3)_2$) were filter sterilized through 0.22-μm membrane filter (Millipore Corporation, Bedford, MA). The plates were incubated at 30°C for 48 h (5). A pure culture of strain NWUAB01 was screened for its ability to grow on different metal concentrations (100 to 1,000 mg/liter) on LB agar plates.

The genomic DNA of strain NWUAB01 was extracted from a pure culture grown on LB agar with the soil microbe DNA Mini Prep DNA extraction kit (Zymo Research, CA) following the kit's protocol. The DNA quality and quantity were determined using a NanoDrop Lite spectrophotometer (Thermo Fischer Scientific, CA). The DNA was sent to Inqaba Biotec, a commercial next-generation sequencing (NGS) service provider in South Africa for 16S rRNA and whole-genome sequencing. The draft genome of *B. cereus* NWUAB01 was obtained with whole-genome shotgun sequencing using an Illumina paired-end library with an average insert size of ~300 bp. We used 50 ng of the DNA sample to prepare the library with the Nextera DNA sample preparation kit (Illumina). The sample was fragmented with an ultrasonication approach (Covaris). The resulting DNA fragments were size selected (300 to 800 bp) with AMPure XP beads. The fragments were then end repaired, and Illumina-specific adapter sequences were ligated to each fragment. The sample was indexed, and a second size selection step was performed. The sample was then quantified with a fluorometric method, diluted to a standard concentration (4 nM), and then sequenced on Illumina's MiSeq platform with a MiSeq v3 (600-cycle) kit as described in the kit protocol.

The raw sequence was processed to obtain high-quality reads with the KBase (6) platform. The read quality was checked with FastQC (v.1.0.4) (7), and the reads were trimmed to remove adapter and low-quality sequences and ambiguous reads using Trimmomatic (v0.36) (8) with default parameters. *De novo* assembly was performed with SPAdes v.3.12.0 (9) with default settings to yield 91 contigs. The draft genome sequence of *B. cereus* NWUAB01 consists of 5,989,415 bp with a G+C content of 35.01%. The contigs have an N_{50} value of 326,240 bp and an L_{50} value of 7, and the largest contig has 644,886 bp. The contigs were annotated with the Rapid Annotations using Subsystems Technology (RAST v 2.0) server (10) and the NCBI Prokaryotic Genome Annotation Pipeline (PGAP v4.7) (11), which identified 6,306 coding genes, 115 RNA, 3 rRNA operons, and 280 pseudogenes.

The genome annotation showed genes responsible for metal resistance, such as the chromate transport protein (ChrA), cobalt-zinc-cadmium resistance protein (CzcD), copper resistance protein (CopC and CopD), arsenic efflux pump protein, cadmium resistance transporter, and cadmium efflux system accessory protein. Analysis of the secondary metabolite biosynthesis gene clusters with antiSMASH v3.0 (12) with default parameters revealed that *B. cereus* NWUAB01 has 44 gene clusters involved in antibiotic and secondary metabolite biosynthesis. These include nonribosomal peptide synthetase (NRPS) gene clusters, lipoproteins, lipopolysaccharides, petrobactin, siderophores, binding proteins, bacillibactin, proteins related to the degradation of toxic compounds, and biofilm secretion genes. No clustered regularly interspaced short palindromic repeat (CRISPR) sequences were identified by CRISPRFinder (13). Further investigation will provide additional information about the mechanism involved in the multimetal resistance role of *B. cereus* NWUAB01 that will facilitate environmental applications.

Data availability. This whole-genome shotgun project is deposited at DDBJ/ENA/GenBank under accession number QNGD00000000, BioProject number PRJNA476495, and BioSample number SAMN09435693. The Sequence Read Archive raw reads are deposited under accession number SRR7647568. The version described in this paper is version QNGD03000000.

REFERENCES

1. Astolfi MCT, de Souza Carvalho EB, de Barros AM, Pinto MV, de Lacerda LB, Nogueira VB, Lopes EF, Astolfi-Filho S. 2018. Draft genome sequence of the novel *Enterobacter cloacae* strain amazonensis, a highly heavy metal-resistant bacterium from a contaminated stream in Amazonas, Brazil. Genome Announc 6:e00450-18. https://doi.org/10.1128/genomeA .00450-18.
2. Rasko DA, Altherr MR, Han CS, Ravel J. 2005. Genomics of the *Bacillus cereus* group of organisms. FEMS Microbiol Rev 29:303–329. https://doi .org/10.1016/j.fmrre.2004.12.005.
3. Singh SK, Tripathi VR, Jain RK, Vikram S, Garg SK. 2010. An antibiotic, heavy metal resistant and halotolerant *Bacillus cereus* SIU1 and its thermoalkaline protease. Microb Cell Fact 9:59. https://doi.org/10.1186/1475-2859-9-59.
4. Lin C-C, Lin H-L. 2005. Remediation of soil contaminated with the heavy metal (Cd^{2+}). J Hazard Mater 122:7–15. https://doi.org/10.1016/j.jhazmat .2005.02.017.
5. Rajkumar M, Freitas H. 2008. Influence of metal resistant-plant growth-promoting bacteria on the growth of *Ricinus communis* in soil contaminated with heavy metals. Chemosphere 71:834–842. https://doi.org/10 .1016/j.chemosphere.2007.11.038.
6. Arkin AP, Stevens RL, Cottingham RW, Maslov S, Henry CS, Dehal P, Ware D, Perez F, Harris NL, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Gunter D, Murphy-Olson D, Chan S, Mamiura RT, Brettin TS, Meyer F, Chivian D, Weston DJ, Glass EM, Davison BH, Kumari S, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M,

Colasanti R, Conrad N, Davis JJ, DeJongh M, Devoid S, Dietrich E, Drake MM, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Mills M, Novichkov P, Oh T, Olsen GJ, Olson B, Parrello B, Pasternak S, Pearson E, Poon SS, Price G, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569. https://doi.org/10.1038/nbt.4163.
7. Babraham Bioinformatics. 2011. FastQC: a quality control tool for high throughput sequence data. Babraham Institute, Cambridge, UK. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
8. 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.
9. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20:714–737. https://doi.org/10.1089/cmb.2013.0084.
10. 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.

11. 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.

12. 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.

13. 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.