



Draft Genome Sequence of *Bacillus* sp. Strain EKM601B (Phylum *Firmicutes*), Living inside the Seeds of *Luffa acutangula* (Chinese Okra)

 Eman M. Khalaf,^{a,b} Manish N. Raizada^a

^aDepartment of Plant Agriculture, University of Guelph, Guelph, Ontario, Canada

^bDepartment of Microbiology and Immunology, Faculty of Pharmacy, Damanhour University, Damanhour, Egypt

ABSTRACT Presented here is the draft genome sequence of *Bacillus* sp. strain EKM601B, which contains 4,199,360 bp in 73 contigs. This candidate endophyte was isolated from surface-sterilized dry seeds of *Luffa acutangula* (Chinese okra) and demonstrated diverse plant-beneficial functions and antagonism against soilborne pathogens *in vitro*.

Seed-associated endophytic bacteria may be vectors for beneficial founder microbes that establish the plant microbiome (1). Many strains of the genus *Bacillus* are marketed as biofertilizers or biocontrol agents. They promote plant growth either directly via nutrient acquisition (2) or indirectly through antibiosis, competition, and induction of host defense responses (3). The seed microbiota of cucurbits (including cucumber, melons, pumpkin, squash, and luffa) were previously cultivated and extensively phenotyped, and *Bacillus* was the dominant genus (4, 5). The endophytic candidate *Bacillus* sp. strain EKM601B (GenBank accession number [KT281323](https://doi.org/10.1128/MRA.00180-20)) was isolated from surface-sterilized seeds of *Luffa acutangula* (Chinese okra) in 2014 (4). The strain exhibited diverse *in vitro* growth-promoting functions, including growth on nitrogen-free medium and secretion of extracellular enzymes (pectinase, protease, and RNase) (4). Furthermore, this candidate endophyte displayed *in vitro* biocontrol activities via emission of volatile organic compounds (VOCs) (acetoin and diacetyl production) known to induce plant defenses and suppression of the soilborne fungal pathogen *Rhizoctonia solani* and the oomycete pathogen *Phytophthora capsici* (5).

Bacterial genomic DNA was isolated from an overnight LB broth culture (37°C, 250 rpm), which had been inoculated from a single colony streaked onto LB agar from the original glycerol stock, by using a DNeasy UltraClean microbial kit (Qiagen product number 12224-50) and then was adjusted to 50 ng/μl. DNA libraries were prepared using a TruSeq DNA Nano library preparation kit (KAPA HyperPrep kit, product number KK8504). The Illumina NovaSeq 6000 platform was used for sequencing, which generated 1,594,540 raw reads with an average length of 150 bp (paired end) with 97-fold coverage, compared to the top genome match in the database (*Bacillus velezensis* strain QST713; GenBank accession number [CP025079.1](https://doi.org/10.1128/MRA.00180-20)) (6), as determined using KmerFinder v3.1 (7) with 99.85% query coverage. Using the EvoCAT (Evogene Clustering and Assembly Toolbox) pipeline, a total of 1,370,507 reads remained after trimming of low-quality sequences using a threshold score of 30. *De novo* assembly using EvoCAT resulted in 73 contigs (minimum scaffold length, 203 bp; maximum length, 644,951 bp; N_{50} , 293,409 bp). The assembled genome is 4,199,360 bp, with a GC content of 47%. Prodigal software (8) was used for protein prediction; predicted proteins were subjected to searches against the NCBI nonredundant protein database using BLASTp (9). Protein domains were identified using InterProScan v5.32-71.0 software (10). Default parameters were used for all software unless otherwise specified.

Citation Khalaf EM, Raizada MN. 2020. Draft genome sequence of *Bacillus* sp. strain EKM601B (phylum *Firmicutes*), living inside the seeds of *Luffa acutangula* (Chinese okra). *Microbiol Resour Announc* 9:e00180-20. <https://doi.org/10.1128/MRA.00180-20>.

Editor David A. Baltrus, University of Arizona

Copyright © 2020 Khalaf and Raizada. 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 Eman M. Khalaf, emankhalaf@pharm.dmu.edu.eg.

Received 10 March 2020

Accepted 27 April 2020

Published 14 May 2020

The *Bacillus* sp. strain EKM601B genome is predicted to encode a minimum of 3,736 proteins. The genome annotation was consistent with the phenotypic traits expressed *in vitro* (4, 5), revealing candidate genes underlying the use of *Bacillus* species as biofertilizers and biocontrol agents (11–14). For example, data mining identified genes required for biological nitrogen fixation (*nifB* gene), secretion of extracellular enzymes that contribute to colonization or antimicrobial activity, such as pectin lyase and diverse proteases (e.g., serine proteases and metalloproteases), antimicrobial peptides (e.g., bacteriocin, which is reported to have broad-spectrum antimicrobial activity against *R. solani* and *P. capsici* [15]), ribonucleases (potential anti-RNA virus activity), butanediol-dehydrogenase-like (acetoin) production, and a wide variety of hydrolytic enzymes (in particular, chitinases [anti-insect and antifungal]). These preliminary findings may help to explain the success of *Bacillus* inoculants in agriculture.

Data availability. This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession number [JAALLJ000000000](https://doi.org/10.1093/nar/gkl842). The version described in this paper is the first version, JAALLJ010000000. Raw Illumina reads are available under SRA accession number [SRR11051678](https://doi.org/10.1093/nar/gkl842).

ACKNOWLEDGMENTS

We thank Ada Viterbo and collaborators at Lavie Bio Ltd. for genome sequencing and bioinformatic support.

This research was supported by grants from the Ontario Ministry of Agriculture, Food, and Rural Affairs, Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Truyens S, Weyens N, Cuypers A, Vangronsveld J. 2015. Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environ Microbiol Rep* 7:40–50. <https://doi.org/10.1111/1758-2229.12181>.
2. Santoyo G, Moreno-Hagelsieb G, del Orozco-Mosqueda MC, Glick BR. 2016. Plant growth-promoting bacterial endophytes. *Microbiol Res* 183: 92–99. <https://doi.org/10.1016/j.micres.2015.11.008>.
3. Compant S, Samad A, Faist H, Sessitsch A. 2019. A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *J Adv Res* 19:29–37. <https://doi.org/10.1016/j.jare.2019.03.004>.
4. Khalaf EM, Raizada MN. 2016. Taxonomic and functional diversity of cultured seed associated microbes of the cucurbit family. *BMC Microbiol* 16:131. <https://doi.org/10.1186/s12866-016-0743-2>.
5. Khalaf EM, Raizada MN. 2018. Bacterial seed endophytes of domesticated cucurbits antagonize fungal and oomycete pathogens including powdery mildew. *Front Microbiol* 9:42. <https://doi.org/10.3389/fmicb.2018.00042>.
6. Pandin C, Le Coq D, Deschamps J, Védie R, Rousseau T, Aymerich S, Briandet R. 2018. Complete genome sequence of *Bacillus velezensis* QST713: a biocontrol agent that protects *Agaricus bisporus* crops against the green mould disease. *J Biotechnol* 278:10–19. <https://doi.org/10.1016/j.jbiotec.2018.04.014>.
7. Deng X, den Bakker HC, Hendriksen RS. 2016. Genomic epidemiology: whole-genome-sequencing-powered surveillance and outbreak investigation of foodborne bacterial pathogens. *Annu Rev Food Sci Technol* 7:353–374. <https://doi.org/10.1146/annurev-food-041715-033259>.
8. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
9. Pruitt KD, Tatusova T, Maglott DR. 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 35:D61–D65. <https://doi.org/10.1093/nar/gkl842>.
10. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Res* 33:W116–W120. <https://doi.org/10.1093/nar/gki442>.
11. Alvarez VM, Von Der Weid I, Seldin L, Santos A. 2006. Influence of growth conditions on the production of extracellular proteolytic enzymes in *Paenibacillus peoriae* NRRL BD-62 and *Paenibacillus polymyxa* SCE2. *Lett Appl Microbiol* 43:625–630. <https://doi.org/10.1111/j.1472-765X.2006.02015.x>.
12. Li S, Yang D, Qiu M, Shao J, Guo R, Shen B, Yin X, Zhang R, Zhang N, Shen Q. 2014. Complete genome sequence of *Paenibacillus polymyxa* SQR-21, a plant growth-promoting rhizobacterium with antifungal activity and rhizosphere colonization ability. *Genome Announc* 2:e00281-14. <https://doi.org/10.1128/genomeA.00281-14>.
13. Ghio S, Martínez Cáceres AI, Talia P, Grasso DH, Campos E. 2015. Draft genome sequence of cellulolytic and xylanolytic *Paenibacillus* sp. A59, isolated from decaying forest soil from Patagonia, Argentina. *Genome Announc* 3:e01233-15. <https://doi.org/10.1128/genomeA.01233-15>.
14. Xie NZ, Li JX, Song LF, Hou JF, Guo L, Du QS, Yu B, Huang RB. 2015. Genome sequence of type strain *Paenibacillus polymyxa* DSM 365, a highly efficient producer of optically active (*R,R*)-2,3-butanediol. *J Biotechnol* 195:72–73. <https://doi.org/10.1016/j.jbiotec.2014.07.441>.
15. Hu HQ, Li XS, He H. 2010. Characterization of an antimicrobial material from a newly isolated *Bacillus amyloliquefaciens* from mangrove for biocontrol of Capsicum bacterial wilt. *Biol Control* 54:359–365. <https://doi.org/10.1016/j.biocontrol.2010.06.015>.