Complete genome sequence of plant growth-promoting *Bacillus stratosphericus* AIMST-CREST02 isolated from bulk soil of a paddy field

Sumitra Sivaprakasam,[1] Nur Arisa Mohd Azim Khan,[2] Tan Yee Fan,[1] Yukgehnaish Kumarasan,[1] Thomas Sicheritz-Pontén,[1,3] Bent Petersen,[1,3] Erneeza Mohd Hata,[2] Ganesan Vadamalai,[2,4] Sivachandran Parimannan,[1,3] Heera Rajandas[1,3]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Here, we present the complete genome of a plant growth-promoting strain, *Bacillus stratosphericus* AIMST-CREST02 isolated from the bulk soil of a high-yielding paddy plot. The genome is 3,840,451 bp in size with a GC content of 41.25%. Annotation predicted the presence of 3,907 coding sequences, including genes involved in auxin biosynthesis regulation and gamma-aminobutyric acid (GABA) metabolism.

The members of the genus *Bacillus* are predominantly known for their plant growth-promoting (PGP) traits and can synthesize various beneficial metabolites even under harsh environmental conditions. This makes them an ideal candidate for biofertilizer development (1, 2). Here, we report the complete genome of a potential PGP *Bacillus stratosphericus* AIMST-CREST02 isolated from the bulk soil sample from a high-yielding paddy plot at Kampung Gajah, Perak, Malaysia (4.1841° N, 100.9389° E) (3–5). We collected 30 grams of submerged bulk soil sample at a depth of 0–20 cm on 1 July 2022.

The strain was initially isolated on a nutrient agar plate incubated at 37°C for 24 hours following the serial dilution of 2.5 grams of paddy bulk soil with sterile distilled water. Upon incubation, a single colony was inoculated and cultivated in 5 mL of fresh Luria-Bertani broth for 24 hours at 37°C. The genomic DNA was then extracted using the GeneJET Genomic DNA Purification Kit (ThermoFisher) following the manufacturer's instructions.

A non-fragmented and non-size-selected DNA library was prepared with SQK-NBD114.24 native barcoding kit. Sequencing was performed with Oxford Nanopore Technology's MinION platform, using an R10.4.1 flow cell. Guppy v6.5.7 (high-accuracy model), a recurrent neural network-based basecaller, was used (6). A total of 277,493 raw reads were generated with a read length N_{50} value of 7,197. The raw reads were trimmed with Porechop v0.2.4 (7) and *de novo* assembled using Flye v2.9.2 (8). The resulting unrotated contig was subsequently polished using Medaka v1.8.1 (9). The publicly available, linked, resource (CP137077) was then annotated with PGAP v6.6 (10). Default parameters were used for all the analyses unless otherwise stated.

Visualization of the assembly graph with Bandage v0.9.0 (11) indicated that the resulting contig was a single circular chromosome (in agreement with Flye's output) of 3,840,451 bp with a GC content of 41.25%, N_{50} length of 3,840,451, and 448× mean coverage. PGAP predicted that the genome encodes 4,017 genes including 3,907 coding sequences, 24 rRNA genes (5S, 16S, and 23S), 81 tRNAs, 5 non-coding RNAs, and 314 pseudogenes. The similar genome finder service on BV-BRC v3.33.16 (12) identified *Bacillus stratosphericus* LAMA 585 (APAS00000000) to be the closest strain to the

assembled genome, and an average nucleotide identity analysis on EzBioCloud (update 2023.08.23) (13) indicated a high sequence similarity of 98.39% between the genomes.

In addition, the presence of several PGP-attributable genes was predicted using the RAST toolkit (RASTtk) (14). The complete genome possessed genes that regulate auxin biosynthesis (anthranilate phosphoribosyltransferase, phosphoribosyl anthranilate isomerase, and tryptophan synthase) and GABA metabolism (gabR). The presence of both GABA and auxin-related genes implies that the isolate will be efficient in alleviating plant stress while enhancing plant growth and development (15). The genome also possessed genes associated with siderophores' production, especially bacillibactin, known to improve iron availability (16, 17). Genes associated with cobalt-zinc-cadmium resistance and volatile organic compound (VOC) metabolism (acetoin and butanediol) beneficial to plants were also found. Therefore, this strain can be used as a potential candidate bacterium for biofertilizer development.

AUTHOR AFFILIATIONS

[1]Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), AIMST University, Bedong, Kedah, Malaysia
[2]Department of Plant Protection, Faculty of Agriculture, Universiti Putra Malaysia, Serdang, Selangor, Malaysia
[3]Center for Evolutionary Hologenomics, Globe Institute, University of Copenhagen, Copenhagen, Denmark
[4]Laboratory of Sustainable Agronomy and Crop Protection, Institute of Plantation Studies, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

AUTHOR ORCIDs

Sivachandran Parimannan https://orcid.org/0000-0002-0101-1858
Heera Rajandas http://orcid.org/0000-0001-5963-3441

AUTHOR CONTRIBUTIONS

Sumitra Sivaprakasam, Formal analysis, Investigation, Methodology, Writing – original draft | Nur Arisa Mohd Azim Khan, Investigation, Methodology | Tan Yee Fan, Investigation, Methodology | Yukgehnaish Kumarasan, Resources, Software | Thomas Sicheritz-Pontén, Resources, Software, Supervision | Bent Petersen, Resources, Software, Writing – review and editing | Erneeza Mohd Hata, Funding acquisition, Investigation, Methodology, Project administration, Resources | Ganesan Vadamalai, Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources

DATA AVAILABILITY

The complete genome sequence of *Bacillus stratosphericus* AIMST-CREST02 has been deposited in NCBI GenBank under the BioProject ID PRJNA1028329, BioSample ID SAMN37821796, SRA ID SRR26661738, and GenBank accession CP137077. The annotation output of RASTtk has been deposited under 10.6084/m9.figshare.25233520.

REFERENCES

1. Hashem A, Tabassum B, Fathi Abd_Allah E. 2019. *Bacillus subtilis*: a plant-growth promoting rhizobacterium that also impacts biotic stress. Saudi J Biol Sci 26:1291–1297. https://doi.org/10.1016/j.sjbs.2019.05.004
2. Tsotetsi T, Nephali L, Malebe M, Tugizimana F. 2022. *Bacillus* for plant growth promotion and stress resilience :what have we learned?. Plants 11:2482. https://doi.org/10.3390/plants11192482
3. Khan MS, Gao J, Zhang M, Chen X, Moe TS, Du Y, Yang F, Xue J, Zhang X. 2020. Isolation and characterization of plant growth-promoting endophytic bacteria *Bacillus stratosphericus* LW-03 from *Lilium wardii*. 3 Biotech 10:305. https://doi.org/10.1007/s13205-020-02294-2
4. Oliva G, Vigliotta G, Terzaghi M, Guarino F, Cicatelli A, Montagnoli A, Castiglione S. 2023. Counteracting action of *Bacillus stratosphericus* and *Staphylococcus succinus* strains against deleterious salt effects on Zea mays L. Front Microbiol 14. https://doi.org/10.3389/fmicb.2023.1171980
5. Durairaj K, Velmurugan P, Park J-H, Chang W-S, Park Y-J, Senthilkumar P, Choi K-M, Lee J-H, Oh B-T. 2017. Potential for plant biocontrol activity of isolated *Pseudomonas aeruginosa* and *Bacillus stratosphericus* strains against bacterial pathogens acting through both induced plant resistance and direct antagonism. FEMS Microbiol Lett 364:fnx225. https://doi.org/10.1093/femsle/fnx225
6. Perešíni P, Boža V, Brejová B, Vinař T. 2021. Nanopore base calling on the edge. Bioinformatics 37:4661–4667. https://doi.org/10.1093/bioinformatics/btab528
7. Wick R. 2018. Porechop: an adapter trimmer for Oxford nanopore reads. https://github.com/rrwick/Porechop.
8. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8
9. Oxford Nanopore Technologies. 2023. Medaka. https://github.com/nanoporetech/medaka.
10. 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
11. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: Interactive visualization of de novo genome assemblies. Bioinformatics 31:3350–3352. https://doi.org/10.1093/bioinformatics/btv383
12. Olson RD, Assaf R, Brettin T, Conrad N, Cucinell C, Davis JJ, Dempsey DM, Dickerman A, Dietrich EM, Kenyon RW, et al. 2023. Introducing the bacterial and viral bioinformatics resource center (BV-BRC): a resource combining PATRIC, IRD and Vipr. Nucleic Acids Res 51:D678–D689. https://doi.org/10.1093/nar/gkac1003
13. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613–1617. https://doi.org/10.1099/ijsem.0.001755
14. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genom 9:75. https://doi.org/10.1186/1471-2164-9-75
15. Olanrewaju OS, Ayangbenro AS, Glick BR, Babalola OO. 2019. Plant health: feedback effect of root exudates-rhizobiome interactions. Appl Microbiol Biotechnol 103:1155–1166. https://doi.org/10.1007/s00253-018-9556-6
16. Kramer J, Özkaya Ö, Kümmerli R. 2020. Bacterial siderophores in community and host interactions. Nat Rev Microbiol 18:152–163. https://doi.org/10.1038/s41579-019-0284-4
17. Singh P, Kumar R, Khan A, Singh A, Srivastava A. 2023. Bacillibactin siderophore induces iron mobilisation responses inside aerobic rice variety through YSL15 transporter. Rhizosphere 27:100724. https://doi.org/10.1016/j.rhisph.2023.100724