

Draft genome of the endophytic *Bacillus velezensis* CNPMS-22 isolated from maize leaves

José Edson Fontes Figueiredo,¹ Mikaelly Sousa Marins,² Felipe Campos Silva,³ Vitória Palhares Ribeiro,² Gisele de Fátima Dias Diniz,² Victor Alef Rodrigues,⁴ Christiane Abreu Oliveira-Paiva,² Ubiraci Gomes de Paula Lana,⁵ Lourenço Vitor Silva Ferreira,⁵ Sylvia Moraes de Sousa Tinoco,⁵ Valter Cruz-Magalhães⁶

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Bacillus velezensis* has been widely used as a biocontrol agent and plant growth promoter in agricultural bio-inputs. The genome of the endophytic bacterial strain CNPMS-22 isolated from maize leaves was sequenced, and the results showed that the strain is a *Bacillus velezensis* species.

KEYWORDS genome analysis, antimicrobial activity, plant growth promoter, endophytes

Bacillus velezensis is a ubiquitous species isolated from diverse niches (1). It exhibits biocontrol activity against fungi, insects, and nematodes and promotes plant growth (2, 3). The endophytic strain CNPMS-22 was isolated in 2007 from maize leaves collected from the experimental field of the Embrapa Maize and Sorghum Research Center (19°26'43.6" S, 44°10'03.0" W) located in Sete Lagoas, Brazil, following a protocol for microbial isolation from plants (4). The bacterium was cryo-preserved at –80°C in a glycerol-based medium and deposited in the Embrapa Collection of Multifunctional and Phytopathogenic Microorganisms of Maize and Sorghum.

For genome sequencing, an aliquot of the glycerol stock was streaked onto a plate containing the Luria-Bertani (LB) agar medium and incubated for 48 h at 28°C. A pure colony recovered from the plate was inoculated into a 5 mL liquid LB medium and incubated for 48 h at 28°C. The genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, USA) and quantified on the Qubit 2.0 fluorometer (Thermo Fisher Scientific Inc. [NYSE: TMO], MA, USA). According to the manufacturer's recommendations, the library was prepared with a Nextera XT kit (Illumina, San Diego, CA, USA). Library quality was checked in GelBot (Loccus, Cotia, SP, Brazil). GoGenetic Company (Curitiba, PR, Brazil) sequenced the genome using the Illumina NextSeq 1000 platform with a P1-600 kit in a paired-end strategy. The Trimmomatic version 0.39.5 software (5) was used for quality control. Sequencing of the CNPMS-22 genome 5,005,400 reads, each 300 bp long, which were used for genome assembly by SPAdes version 3.12.0 (6). The assembly quality was assessed using Quast version 5.0.2 (7), and the completeness was evaluated using Benchmarking Universal Single-Copy Orthologs (version 5.3.1) (8) with the *bacteria_odb10* lineage data set ($n = 124$ bacterial orthologs). The NCBI Prokaryotic Genome Annotation Pipeline version 6.8 (9) performed the draft genome annotation.

The CNPMS-22 genome sequencing generated 5,005,400 paired-end reads, totaling ≈1.5 gigabytes of data and a theoretical coverage ≈373×. After quality filtering, the assembly generated 123 contigs with an N_{50} value of 258,002 and a final coverage genome equal to 8.0× across retained contigs. The genome has a length of 4,040,390 bp and 4,038 genes for 3,813 proteins, 84 tRNAs, 14 rRNAs, 5 ncRNAs, 1 tmRNA, and 46.5% G + C content. To confirm the taxonomic position of CNPMS-22, we used the fastANI

Editor Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

Address correspondence to José Edson Fontes Figueiredo, jose.edson@embrapa.br.

The authors declare no conflict of interest.

Received 27 November 2024

Accepted 8 April 2025

Published 7 May 2025

Copyright © 2025 Figueiredo 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/).

program (10) and the Type Strain Genome Server (version 39.1) (11) for digital DNA-DNA hybridization (dDDH). The average nucleotide identity (ANI) value and taxonomic index dDDH were calculated concerning the type strain *B. velezensis* FZB42 (GenBank accession number [NC_009725.2](#)) (12). The ANI values and DDH index were 98.67% and 96.4%, respectively, which confirmed that CNPMS-22 is phylogenetically related to the type strain FZB42. Default parameters were used for all software.

Genome analysis and manual genome mining identified the three gene clusters predicted to synthesize iturin, surfactin, and fengycin, corroborating our previous ultra-performance liquid chromatography-mass spectrometry studies with CNPMS-22 (13). We also found that CNPMS-22 harbors other non-ribosomal, ribosomal, and polyketide genes predicted to encode for molecules with antifungal activity.

ACKNOWLEDGMENTS

The authors thank Brazilian Agricultural Research Enterprise—Embrapa and FAPEMIG for financial support.

AUTHOR AFFILIATIONS

¹Biochemistry Molecular Laboratory, Embrapa Maize and Sorghum, Sete Lagoas, Minas Gerais, Brazil

²Soil Microbiology Laboratory, Embrapa Maize and Sorghum, Sete Lagoas, Minas Gerais, Brazil

³Microbiology Laboratory, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil

⁴Soil Microbiology Laboratory, Universidade Federal de São João del Rei-UFSJ-CSL, Sete Lagoas, Minas Gerais, Brazil

⁵Núcleo de Biologia Aplicada, Embrapa Maize and Sorghum, Sete Lagoas, Minas Gerais, Brazil

⁶Microbial Ecology Laboratory, Embrapa Maize and Sorghum, Sete Lagoas, Minas Gerais, Brazil

AUTHOR ORCIDs

José Edson Fontes Figueiredo  <http://orcid.org/0000-0002-9542-2315>

Mikaely Sousa Marins  <http://orcid.org/0000-0002-3508-8204>

Felipe Campos Silva  <https://orcid.org/0009-0001-1301-3519>

Vitória Palhares Ribeiro  <http://orcid.org/0000-0003-1675-0669>

Gisele de Fátima Dias Diniz  <http://orcid.org/0000-0001-7018-6574>

Victor Alef Rodrigues  <http://orcid.org/0009-0001-4245-8664>

Christiane Abreu Oliveira-Paiva  <http://orcid.org/0000-0002-1874-2489>

Ubiraci Gomes de Paula Lana  <http://orcid.org/0000-0003-3964-8966>

Lourenço Vitor Silva Ferreira  <http://orcid.org/0009-0003-8797-5213>

Sylvia Morais de Sousa Tinoco  <http://orcid.org/0000-0002-8030-1385>

Valter Cruz-Magalhães  <http://orcid.org/0000-0001-5552-1442>

AUTHOR CONTRIBUTIONS

José Edson Fontes Figueiredo, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Mikaely Sousa Marins, Formal analysis, Methodology | Felipe Campos Silva, Formal analysis, Methodology | Vitória Palhares Ribeiro, Conceptualization, Formal analysis | Gisele de Fátima Dias Diniz, Conceptualization, Methodology, Validation | Victor Alef Rodrigues, Investigation, Methodology | Christiane Abreu Oliveira-Paiva, Funding acquisition, Resources, Visualization, Writing – review and editing | Ubiraci Gomes de Paula Lana, Methodology | Lourenço Vitor Silva Ferreira, Data curation, Formal analysis, Methodology, Validation | Sylvia Morais de Sousa Tinoco, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Validation |

Valter Cruz-Magalhães, Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft

DATA AVAILABILITY

The draft genome sequence of *B. velezensis* strain CNPMS-22 has been deposited in GenBank BioProject [PRJNA1108656](https://doi.org/10.1093/bioinformatics/bty266), BioSample [SAMN43568027](https://doi.org/10.1093/bioinformatics/bty266), accession number [JBHGBV000000000.1](https://doi.org/10.1093/bioinformatics/bty266), contig-level assemblies [JBHGBV01](https://doi.org/10.1093/bioinformatics/bty266), NCBI RefSeq assembly [GCF_041949825.1](https://doi.org/10.1093/bioinformatics/bty266), and Sequence Read Archive (SRA) [SRR31156977](https://doi.org/10.1093/bioinformatics/bty266).

REFERENCES

- Alenezi FN, Slama HB, Bouket AC, Cherif-Silini H, Silini A, Luptakova L, Nowakowska JA, Oszako T, Belbahri L. 2021. *Bacillus velezensis*: a treasure house of bioactive compounds of medicinal, biocontrol and environmental importance. *Forests* 12:1714. <https://doi.org/10.3390/f12121714>
- Fan B, Wang C, Song X, Ding X, Wu L, Wu H, Gao X, Borris R. 2018. *Bacillus velezensis* FZB42 in 2018: the gram-positive model strain for plant growth promotion and biocontrol. *Front Microbiol* 9:2491. <https://doi.org/10.3389/fmicb.2018.02491>
- Figueiredo JEF, Diniz G de FD, Marins MS, Silva FC, Ribeiro VP, Lanza FE, de Oliveira-Paiva CA, Cruz-Magalhães V. 2025. *Bacillus velezensis* CNPMS-22 as biocontrol agent of pathogenic fungi and plant growth promoter. *Front Microbiol* 16:1522136. <https://doi.org/10.3389/fmicb.2025.1522136>
- Figueiredo JEF, Gomes EA, Guimarães CT, de Paula Lana UG, Teixeira MA, Lima GVC, Bressan W. 2009. Molecular analysis of endophytic bacteria from the genus *Bacillus* isolated from tropical maize (*Zea mays* L.). *Braz J Microbiol* 40:522–534. <https://doi.org/10.1590/S1517-838220090003000014>
- 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>
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>
- Mikheenko A, Pribelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34:i142–i150. <https://doi.org/10.1093/bioinformatics/bty266>
- Manni M, Berkeley MR, Seppely M, Zdobnov EM. 2021. BUSCO: assessing genomic data quality and beyond. *Curr Protoc* 1:e323. <https://doi.org/10.1002/cpz1.323>
- 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>
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>
- Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>
- Fritze D. 2004. Taxonomy of the genus *Bacillus* and related genera: the aerobic endospore-forming bacteria. *Phytopathology* 94:1245–1248. <https://doi.org/10.1094/PHYTO.2004.94.11.1245>
- de Souza CG, Martins FICC, Zocolo GJ, Figueiredo JEF, Canuto KM, de Brito ES. 2018. Simultaneous quantification of lipopeptide isoforms by UPLC-MS in the fermentation broth from *Bacillus subtilis* CNPMS22. *Anal Bioanal Chem* 410:6827–6836. <https://doi.org/10.1007/s00216-018-1281-6>