

Draft genome sequences of representative *Paenibacillus polymyxa*, *Bacillus cereus*, *Fictibacillus* sp., and *Brevibacillus agri* strains isolated from Amazonian dark earth

Thierry Alexandre Pellegrinetti,¹ Izadora De Cássia Mesquita da Cunha,^{1,2} Miriam Gonçalves de Chaves,¹ Anderson Santos de Freitas,¹ Ana Vitória Reina da Silva,¹ Siu Mui Tsai,¹ Lucas William Mendes¹

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT Here, we report 10 distinct bacterial genomes from Amazonian dark earths, including six identified as *Paenibacillus polymyxa*, while the remaining four were unique representatives of *Paenibacillus vini*, *Bacillus cereus*, *Brevibacillus agri*, and *Fictibacillus* sp., respectively. Each strain exhibited antagonistic activity against *Fusarium oxysporum*, underscoring their potential as sustainable agriculture resources.

KEYWORDS plant-microbe interaction, rhizosphere-inhabiting microbes, *Fusarium*, terra preta

Amazonian dark earths (ADEs), a fertile soil type formed by the management practices of pre-Columbian indigenous near Amazonian rivers over two millennia ago (1), is known for promoting plant growth and nutrient retention. Its diverse microbial communities play key roles in nutrient cycling and soil fertility, offering the potential for agricultural and ecological advancements (2). Microbial isolation from ADE allows for a detailed study of the soil's distinctive chemical and biological properties. This understanding paves the way for insights into the ecological roles of resident microorganisms and their potential biotechnological applications in areas such as sustainable agriculture, waste management, and biofuel production (3).

The strains presented here were previously isolated in 2016 from an experiment using ADE soils and two different cultivars of common bean, exhibiting contrasting levels of resistance to the soil-borne pathogen *Fusarium oxysporum*. The experiment was conducted at the Center for Nuclear Energy in Agriculture of the University of São Paulo (22° 42' 27.60" S, 47° 38' 41.17" W) (4). Plants were collected at the R1 development stage (early flowering), and the roots were shaken to remove loosely adhering soil. The firmly attached soil was collected with sterile brushes and considered to be the rhizosphere soil. For microbial isolation, 1 g of rhizosphere soil was mixed with 9 mL of saline solution (8.5 g L⁻¹ NaCl). Serial dilutions (10⁻¹ to 10⁻⁶) were performed and then transferred onto KING medium plates (5). After incubating at 25°C for 48 h, colonies were isolated using the streak plate method. Total DNA from isolates was extracted using the DNeasy Ultraclean Microbial Kit (Qiagen Mobio, Hilden, Germany) following the manufacturer's instructions.

Whole-genome sequencing was performed using the Illumina HiSeq 2500 platform. Genomic DNA was used for paired-end libraries (2 × 150 bp) with NEBNext Fast DNA Fragmentation and Library Preparation Kit (New England Biolabs Inc., MA, USA). The number of library reads is described in Table 1. We visualized the quality of raw sequences obtained from Illumina HiSeq sequencing by using FastQC v0.12.1 (6). Given that all raw reads exhibited a quality score exceeding Q30, trimming was not performed. Then, the sequences underwent *de novo* assembly using SPAdes v3.15.5

Editor Simon Roux, DOE Joint Genome Institute, Berkeley, California, USA

Address correspondence to Lucas William Mendes, lwmdes@cena.usp.br.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 7 July 2023

Accepted 26 July 2023

Published 9 October 2023

Copyright © 2023 Pellegrinetti 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/).

TABLE 1 Assembly properties of 10 strains isolated from Amazonian dark earth, including contig statistics, genome size, GC content, and predicted gene count

Isolate	<i>P. vini</i>		<i>P. polymyxa</i>		<i>Fictibacillus</i> sp.		<i>B. agri</i>		<i>P. polymyxa</i>		<i>B. cereus</i>		<i>P. polymyxa</i>		<i>P. polymyxa</i>	
	CENA-BCM001	CENA-BCM002	CENA-BCM003	CENA-BCM004	CENA-BCM005	CENA-BCM006	CENA-BCM007	CENA-BCM008	CENA-BCM009	CENA-BCM010	CENA-BCM011	CENA-BCM012	CENA-BCM013	CENA-BCM014	CENA-BCM015	CENA-BCM016
Library size (million of reads)	20.9	23.2	25.7	24	26.8	21.6	19.6	24.4	26.7	22.3						
Number of contigs	65	57	32	16	111	34	22	34	31	30						
Largest contig (bp)	343,800	689,511	1,524,845	1,557,022	1,557,022	371,701	1,100,025	1,517,151	1,524,846	1,445,065						
Total length (bp)	5,712,199	5,635,190	5,849,366	4,967,627	5,280,964	5,280,964	5,569,359	5,726,433	5,840,294	5,711,641						
N50 (bp)	188,024	202,973	686,293	1,390,346	1,390,346	79,882	668,640	1,443,580	686,293	690,213						
N75 (bp)	101,676	128,558	230,621	407,875	407,875	47,434	362,608	207,366	186,286	206,180						
GC (%)	48.88	45.28	45.36	43.65	43.65	53.62	34.96	45.40	45.36	45.38						
Genes	5,274	4,975	5,169	5,118	5,255	5,446	5,635	5,075	5,162	5,011						
Protein coding	5,159	4,913	5,044	4,882	5,060	5,310	5,459	4,950	5,033	4,907						
Non-coding	72	22	75	64	72	75	58	79	79	61						
Coverage (x)	1082.86	1082.86	1305.91	1437.35	1471.72	1041.85	1047.66	1269.73	1359.57	1160.26						
SRA identifiers	SRX20866903	SRX20866904	SRX20866905	SRX20866906	SRX20866907	SRX20866908	SRX20866909	SRX20866910	SRX20866911	SRX20866912						
GenBank identifiers	GCA_030412165.1	GCA_030412275.1	GCA_030412235.1	GCA_030412265.1	GCA_030412255.1	GCA_030412175.1	GCA_030412285.1	GCA_030412135.1	GCA_030412145.1	GCA_030412155.1						

(7). We employed QUAST v4.4 (8) to measure the assembled genomes' quality metrics. Paired-end reads were aligned to the assemblies using the Burrows-Wheeler Aligner tool v0.7.17 (9). The average sequencing coverage was then calculated from the resulting BAM (Binary Alignment Map) file using SAMtools v1.15 (10). Taxonomic classification was performed with GTDB-tk v2.3.2 (11). The genome annotation was performed by NCBI Prokaryotic Genome Annotation Process v6.5 (12). An additional functional gene annotation was carried out using RAST-tk 1.3.0 (13). Default parameters were used for all software. Genome assembly statistics and information about the strains are summarized in Table 1.

ACKNOWLEDGMENTS

This study was supported by grants from Fundação de Amparo à pesquisa do Estado de São Paulo (FAPESP 2014/03217-3 and 2015/00251-9), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES 88887.185941/2018-00), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 408191/2018-0), and PRPI-USP (22.1.08498.01.0). L.W.M. also thanks CNPq for the Productivity Research Grant (CNPq 307670/2021-0).

All authors meet the MRA authorship criteria. T.A.P., M.G.C., and L.W.M. were responsible for the organization and coordination of the analysis. L.W.M. was the chief investigator, and T.A.P. was responsible for the data analysis. All authors contributed to the writing of the final manuscript.

AUTHOR AFFILIATIONS

¹Cell and Molecular Biology Laboratory, Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, Brazil

²Luiz de Queiroz College of Agriculture (ESALQ), University of São Paulo, Piracicaba, Brazil

AUTHOR ORCIDs

Siu Mui Tsai  <http://orcid.org/0000-0002-3733-6312>

Lucas William Mendes  <http://orcid.org/0000-0003-0980-7006>

FUNDING

Funder	Grant(s)	Author(s)
Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)	2014/03217-3, 2015/00251-9	Lucas William Mendes
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)	88887.185941/2018-00	Lucas William Mendes
Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)	408191/2018-0, 307670/2021-0	Lucas William Mendes
PRPI-USP	22.1.08498.01.0	Thierry Alexandre Pellegrinetti

AUTHOR CONTRIBUTIONS

Thierry Alexandre Pellegrinetti, Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review and editing | Izadora De Cássia Mesquita da Cunha, Investigation, Methodology | Miriam Gonçalves de Chaves, Investigation, Methodology | Anderson Santos de Freitas, Formal analysis | Ana Vitória Reina da Silva, Investigation | Siu Mui Tsai, Funding acquisition, Supervision.

DATA AVAILABILITY

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the BioProject accession number [PRJNA988090](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA988090). The genome sequences are deposited under BioSample accession numbers [SAMN35999008](https://www.ncbi.nlm.nih.gov/biosample/SAMN35999008) to [SAMN35999017](https://www.ncbi.nlm.nih.gov/biosample/SAMN35999017) and raw sequences under [SAMN36017836](https://www.ncbi.nlm.nih.gov/biosample/SAMN36017836) to [SAMN36017845](https://www.ncbi.nlm.nih.gov/biosample/SAMN36017845). The SRA and GenBank identifiers are described in Table 1. RASTtk annotations are available in zenodo doi: [10.5281/zenodo.8087953](https://doi.org/10.5281/zenodo.8087953).

REFERENCES

1. Glaser B, Birk JJ. 2012. State of the scientific knowledge on properties and genesis of Anthropogenic dark Earths in Central Amazonia (*terra preta de Índio*). *Geochimica et Cosmochimica Acta* 82:39–51. <https://doi.org/10.1016/j.gca.2010.11.029>
2. Freitas AS de, Zagatto LFG, Rocha GS, Muchalak F, Silva S dos S, Muniz AW, Hanada RE, Tsai SM. 2023. Amazonian dark Earths enhance the establishment of tree species in forest ecological restoration. *Front Soil Sci* 3. <https://doi.org/10.3389/fsoil.2023.1161627>
3. Brossi MJ de L, Mendes LW, Germano MG, Lima AB, Tsai SM, Ahmed N. 2014. Assessment of bacterial *bph* gene in Amazon Dark Earth and their adjacent soils. *PLoS ONE* 9:e99597. <https://doi.org/10.1371/journal.pone.0099597>
4. Mendes LW, de Chaves MG, Fonseca M de C, Mendes R, Raaijmakers JM, Tsai SM. 2019. Resistance breeding of common bean shapes the physiology of the rhizosphere microbiome. *Front Microbiol* 10:2252. <https://doi.org/10.3389/fmicb.2019.02252>
5. KING EO, WARD MK, RANEY DE. 1954. Two simple media for the demonstration of pyocyanin and fluorescin. *J Lab Clin Med* 44:301–307.
6. Andrews S. 2010. Fastqc: A quality control tool for high throughput sequence data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
7. 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>
8. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
9. Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
10. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map (SAM) format and samtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
11. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH, Hancock J. 2020. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>
12. 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>
13. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA 3rd, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. *RASTtk*: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>