Discover Applied Sciences

Brief Communication

Genome sequencing of *Bacillus cereus* isolate EB-40: a plant growth-promoting bacterium

Débora Francine Gomes Silva Pereira[1] · Pedro Marcus Pereira Vidigal[2] · Samuel A. Santos[3] · Silvia Nietsche[4] · Adelica Aparecida Xavier[5] · Marlon Cristian Toledo Pereira[5]

Received: 23 May 2024 / Accepted: 17 October 2024
Published online: 25 October 2024

Abstract
Bacillus sp. isolate EB-40 was characterized in 'Prata Anã' banana (*Musa* sp.) plants as an endophyte capable of colonizing both the inter- and intracellular spaces of roots, nitrogen fixation, phosphate solubilization, in vitro synthesis of indole-3-acetic acid, and promotion of the development of micropropagated banana seedlings. Here, we report the whole-genome sequence of *Bacillus* sp. isolate EB-40 and its taxonomic assignment. Its genome is composed of one chromosome and three plasmids. The circular double-stranded DNA chromosome (5,613,235 base pairs (bp)) has a GC content of 35.3% and 5462 genes. The three plasmids have a total length of 237,685 bp with 201 genes. Comparative genomics revealed that its genome shares similarity indices with *Bacillus cereus* genomes above the thresholds recommended for species assignment.

Article Highlights

- B. cereus EB-40 is a novel bacterium with phosphorus-solubilizing activity, nitrogen fixation ability and IAA production ability.
- The genome of B. cereus EB-40 consists of a circular double-stranded chromosome and three plasmids.
- The complete genome sequence of strain EB-40 provides a genetic basis for multifunctional biofertilizers.

Keywords *Bacillus* · *Musa* · Plant growth · Whole-genome sequence

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42452-024-06284-4.

✉ Silvia Nietsche, snietsche@ufmg.br | [1]Instituto Federal Baiano, Bom Jesus da Lapa, BA 47600-000, Brazil. [2]Núcleo de Análise de Biomoléculas (NuBioMol), Campus da UFV, Universidade Federal de Viçosa (UFV), Viçosa, MG 36570-075, Brazil. [3]Departamento de Fitopatologia, Universidade Federal de Viçosa (UFV), Viçosa, MG 36570-075, Brazil. [4]Instituto de Ciências Agrárias, Universidade Federal de Minas Gerais, Montes Claros, MG 39404-547, Brazil. [5]Departamento de Ciências Agrárias, Universidade Estadual de Montes Claros, Janaúba, MG 39448-581, Brazil.

1 Introduction

Plant growth-promoting bacteria (PGPB) belong to a group of beneficial and heterogeneous microorganisms that can be found in the rhizosphere, on the surface of or associated with the root, and are capable of increasing plant growth and protecting plants from diseases and abiotic stresses [1]. Among the various PGPB, some species of the *Bacillus* genus stand out [2]. They are considered good options for the formulation of commercial products [3].

Bacillus species can directly or indirectly stimulate plant growth by providing plants with resources or nutrients and modulating growth regulator levels in agriculturally significant crops, and these effects improve the productivity of plants and their content of biologically active substances [4, 5]. *Bacillus* sp. EB-40, which was isolated by Souza et al. (2013) [6] from 'Prata Anã' banana (*Musa* spp.) roots, has been studied for over a decade by our group. Our research has demonstrated its ability to colonize both the inter- and intracellular spaces of banana roots [7], fix nitrogen, solubilize phosphate [8], and promote increases in the length and diameter of pseudostems, fresh mass, and dry mass in micropropagated seedlings of 'Prata Anã' banana during the acclimatization period [9]. Furthermore, when used in combination with other bacteria (*Bacillus pumilus* EB-51 and *Lysinibacillus* sp. EB-53), *Bacillus* sp. EB-40 led to a 174% increase in root dry weight and facilitated increased accumulation of macroelements in the aerial parts of micropropagated banana seedlings during the acclimatization period [8]. Although the mechanisms of action of *Bacillus* sp. EB-40 in promoting plant growth have been extensively studied, its genetic background and taxonomy remain unknown.

The development of sequencing methodologies has revolutionized microbial genomics, and approaches such as hybrid assembly have emerged as solutions for obtaining complete and accurate bacterial genomes [10]. The quality of genome assembly has been a concern in microbial genomics, as it directly impacts the accuracy of gene annotation and can hinder comparative genomic analyses [11]. Among the 7,017 genomes of *Bacillus* species available in the NCBI Microbial Genomes database (https://www.ncbi.nlm.nih.gov/genome/microbes/; accessed on 2023/03/24), 1,687 are complete genomes.

The active principles of PGPB are mediated by enzymes and secondary metabolites, and genome mining offers the opportunity to scrutinize the whole genome of a PGPB strain for genes encoding beneficial enzymes. Whole-genome analyses combined with biochemical assays offer the advantage that the full arsenal of a PGPR strain can be unveiled and investigated [12]. Here, we report the whole-genome sequence of *Bacillus* sp. EB-40 and its taxonomic assignment on the basis of its genomic features.

2 Materials and methods

2.1 Genome sequencing, assembly, and annotation

Bacillus sp. EB-40 was previously isolated from the internal tissues of banana roots by Souza et al. (2013) [6]. We multiplied the isolate EB-40 by streaking on a Petri dish containing Luria Bertani medium (2%) and incubating it at 37 °C for 24 h. Then, we selected one isolated colony and cultured it in 10 ml of Luria Bertani medium (2%) for 24 h at 37 °C and 200 RPM. We subsequently extracted the genomic DNA via the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) and assessed its quality and quantity via a NanoDrop™ fluorometer and a Qubit™ fluorometer (Thermo Fischer Scientific, Waltham, MA, USA), respectively. A hybrid approach combining Illumina NextSeq 550 short paired reads (2 × 150 nt) and PacBio RSII long reads was used to determine the complete genome of *Bacillus* sp. EB-40. The quality of the short-read sequencing data was assessed via FASTQC version 0.11.9 (https://github.com/s-andrews/FastQC). Adapter sequences were detected and removed from the sequencing data via the autodetection setting of TrimGalore version 0.6.7 [13]. The paired reads were subsequently trimmed for quality and filtered for length via Trimmomatic version 0.39 [14] with the following parameters: HEADCROP:25, CROP:125, SLIDINGWINDOW:4:20, and MINLEN:100. The long reads were corrected and trimmed by LoRDEC version 0.3 [15] using the processed short reads, and Seqkit filtered the long reads with lengths equal to or greater than 1,000 nt. Unicycler version 0.5.0 [16] was used with both long and short reads to execute the de novo assembly of the genome via the normal method, and BBMap version 38.76 (https://sourceforge.net/projects/bbmap) was used to calculate the sequencing coverage. The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [17–19] annotated the genome during the processing

of its submission to the GenBank database (https://www.ncbi.nlm.nih.gov/genbank), and BUSCO version 5.4.2 [20] was used to assess the completeness of the genome using the bacillales_odb10 database. The Rapid Annotation Subsystem Technology (RAST) server (https://rast.nmpdr.org/) [21] classified the predicted genes into subsystems of the SEED database [22], and the RPSBLAST tool of BLAST version 2.13.0 [23] classified the proteins according to the functional categories of the Clusters of Orthologous Genes (COG) database [24], considering an E-value threshold of $1e^{-10}$ to select the significant alignments. In addition, the KEGG Automatic Annotation Server (KASS) [25] was used to classify the proteins according to the KEGG Orthology (KO) system to allow metabolic pathway mapping.

3 Results and discussion

3.1 Genome sequencing

We sequenced the *Bacillus* sp. EB-40 genome with 745X coverage and assembled it into a single chromosome totaling 5,613,235 bp with an average GC content of 35.3%. The genome also harbors three plasmids with sizes of 215,503, 7,710, and 14,472 bp, each with GC contents of 32.6, 33.8, and 34.2%, respectively (Table 1**,** Fig. 1). The chromosome has 5,462 genes predicted by the NCBI PGAP, including 5,201 coding DNA sequences (CDSs), 106 tRNA genes, 42 rRNA genes, one tmRNA, four ncRNAs, and 108 pseudogenes that encode nonfunctional proteins. Additionally, the three plasmids possess a set of 185 CDSs and 16 pseudogenes. The BUSCO analysis estimated a completeness of 99.8% for the genome of *Bacillus* sp. EB-40 among the 450 BUSCO groups expected for the order Bacillales, identifying 444 complete and single-copy BUSCOs, five complete and duplicate BUSCOs, and only one missing BUSCO.

The species is within the *B. cereus* group, which is composed of rod-shaped, aerobic or facultatively anaerobic spore-forming gram-positive bacteria widely distributed in natural environments [26]. Several works have demonstrated the role of several species of *Bacillus cereus* as bacteria that promote plant growth [26, 27].

The RAST server classified 2,225 CDSs (41.31%) into 27 functional categories of the SEED database, and the most frequent were amino acids and derivatives (257 CDSs); cofactors, vitamins, prosthetic groups, and pigments (202 CDSs); and carbohydrates (191 CDSs). The COG database allowed the classification of 3,429 CDSs (63.67%) into 21 functional categories, and [E] amino transport and metabolism (339 CDSs); [K] transcription (304 CDSs); and [J] translation, ribosomal structure, and metabolism (266 CDSs) were the most frequent among those with known functions. Among the 1,702 CDSs (31.60%) that encode unclassified proteins, 562 CDSs encode hypothetical proteins. Table S1 summarizes the functional classification of the predicted proteins.

Table 1 List of features associated with the *B. cereus* isolate EB-40 genome

Item	Features
Project name	Genome sequencing and assembly of *Bacillus cereus* isolate EB-40
Sample name	EB-40
Taxonomy ID Accession number	1396 [*Bacillus cereus* species] CP115717
Latitude and longitude	14° 13′ 12.0" S and 42° 46′ 48.0" W
Geographic location	Brazil, Bahia, Guanambi
Collection date	2010
Environment	banana roots (*Musa* spp.)
Sequencing method	Illumina NextSeq 550 and PacBio RSII
Number of replicons	4
Reference for biomaterial	Not applicable
Isolation and growth condition	Luria Bertani medium at 37 °C for 24 h
Assembly quality	Finished genome
Assembly software	Unicycler version 0.5.0
Number of contigs	4

All features are in accordance with the recommendations of the minimum information about the bacterial and archaeal genome sequences (MIGS-BA) (https://www.gensc.org/pages/standards/checklists.html)

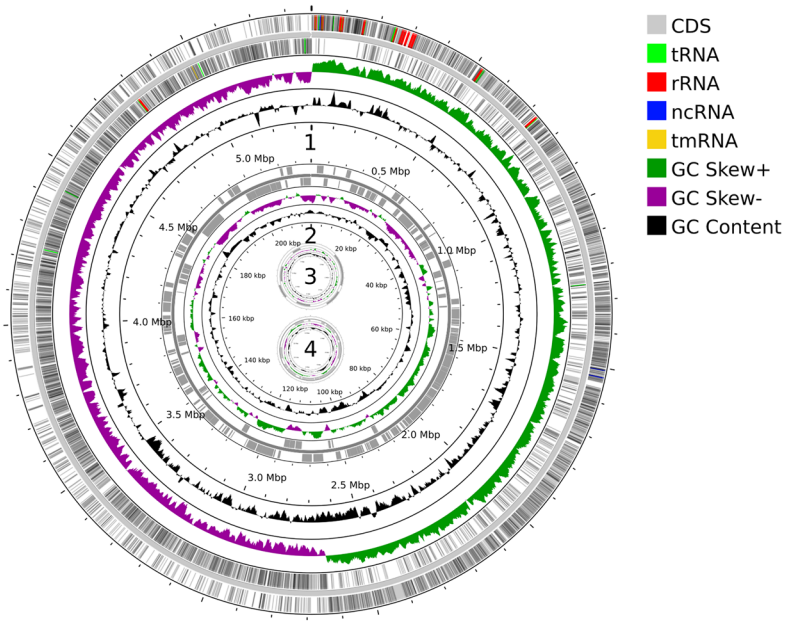

Fig. 1 Genome map of the *Bacillus cereus isolate* EB-40. The genome contains one chromosome (ID 1) and three plasmids (IDs 2, 3, and 4). Each representation harbors four circles from outside to inside: I. The annotated genes in the heavy strand and light strand (inside); II. The average GC skew; III. The GC content; IV. The DNA coordinates. The colors indicate the categories of the predicted genes: coding DNA sequences (CDS), tRNA, rRNA, ncRNA, and tmRNA. The Proksee web server was used to generate the map (https://proksee.ca; [22])

ID	Accession	Topology	Size (nt)	% GC	Avg. Coverage	Genes
1	CP115717	circular(complete)	5,375,280	0.353	745	5462
2	CP115718	circular(complete)	215,503	0.326	446	179
3	CP115719	circular(complete)	7,710	0,338	7,866	8
4	CP115720	linear(partial)	14,742	0.342	265	14

Considering the aforementioned phenotypic characteristics of the *Bacillus* sp. isolate EB-40, we conducted a comprehensive analysis of its genome to identify genes associated with phosphate metabolism, the production of indole-3-acetic acid (IAA or 3-IAA), and volatile organic compounds (VOCs).

Some soil endophytic microorganisms can colonize the rhizosphere of plants and produce extracellular phytase enzymes that can act as plant growth-promoting rhizobacteria by increasing the availability of phytate phosphate [23]. In the genome of *Bacillus* sp. EB-40, we identified the following enzymes that catabolize phytate into forms assimilable by plants: l-myo-inositol-1,4-monophosphatase (*locus* PF061_19825), phospholipase C (*locus* PF061_03420), 1-phosphatidylinositol phosphodiesterase (*locus* PF061_17290), malonate-semialdehyde dehydrogenase (*locus* PF061_11600), and cyclo-inositol 2-dehydrogenase ($NADP^+$) (*locus* PF061_17765). The identification of these enzymes corroborates the results obtained in vitro and in vivo by Andrade et al. (2014) [24], who demonstrated phosphate solubilization, and Souza et al. (2017) [5], who investigated the growth of banana seedlings inoculated with *Bacillus* sp. isolate EB-40 during the acclimatization period in greenhouse experiments.

Bacillus sp. EB-40 also produces auxin in media with and without the addition of tryptophan. Considering tryptophan-dependent pathways, endophytic microorganisms can synthesize indole-3-acetic acid (IAA) via different pathway routes classified based on involved intermediates, such as indole-3-acetamide (IAM), indole-3-pyruvic acid (IPA), indole-3-acetonitrile (IAN), tryptamine (TAM), and tryptophan side-chain oxidase (TSO) pathways [25, 26]. In the genome of the EB-40 isolate, we identified enzymes associated with the TAM, IAM, and TSO pathways. Among the genes involved in tryptophan metabolism, we identified *trp*A (tryptophan synthase subunit alpha; *locus* PF061_06305), trpB (tryptophan synthase subunit beta; *locus* PF061_06300), *trp*C (indole-3-glycerol phosphate synthase; *locus* PF061_06290), *trp*D (anthranilate phosphoribosyltransferase 2; *locus* PF061_06285), *trp*E (anthranilate synthase component I; *loci* PF061_00390 and PF061_06275), *trp*F (phosphoribosylanthranilate isomerase; loci PF061_06295), and trpG (anthranilate synthase component II; *loci* PF061_00395 and PF061_06280).

Volatile organic compounds (VOCs) have the potential to control plant pathogens, stimulate plant growth, and induce systemic disease resistance [27]. Acetoin and 2,3-butanediol are VOCs produced by plant growth-promoting bacteria and are synthesized through the condensation of two pyruvate molecules into acetolactate. Acetolactate synthase catalyzes the production of acetolactate, which is decarboxylated into acetoin by acetolactate decarboxylase. Finally, the reduction of acetoin by the enzyme acetoin reductase produces 2,3-butanediol [28]. The genome of *Bacillus* sp. isolate EB-40 contains all the genes necessary for the production of acetoin and 2,3-butanediol: (R,R)-butanediol dehydrogenase (*locus*

Table 2 Taxonomic assignment of *Bacillus* sp. EB-40

GenBank ID	Species	dDDH [P(DDH ≥ 70%)]	ANIb [Coverage]	ANIm [Coverage]
NZ_CP017060	*Bacillus cereus*	70.80 [79.67]	96.30 [87.81]	96.70 [90.77]
NZ_CP049019	*Bacillus tropicus*	46.30 [10.46]	91.56 [79.43]	92.28 [83.03]
NZ_CP101135	*Bacillus paranthracis*	45.60 [9.14]	91.29 [77.22]	92.06 [80.60]
NZ_NWUW00000000	*Bacillus fungorum*	45.40 [8.89]	90.87 [73.53]	92.06 [76.61]
NC_007530	*Bacillus anthracis*	45.10 [8.33]	91.14 [78.59]	91.94 [81.87]
NZ_CP086328	*Bacillus pacificus*	45.10 [8.32]	91.05 [76.39]	91.92 [79.90]
NZ_CP032365	*Bacillus wiedmannii*	44.70 [7.69]	91.05 [79.10]	91.83 [82.59]
NZ_CP064875	*Bacillus toyonensis*	44.60 [7.60]	91.17 [80.77]	91.82 [84.97]
NZ_CP128152	*Bacillus albus*	44.40 [7.24]	91.00 [80.29]	91.71 [84.12]
NZ_CP040336	*Bacillus luti*	43.60 [6.18]	90.57 [76.14]	91.56 [79.42]

The comparative genomic analysis of the isolate EB-40 genome with the reference prokaryotic genomes of the NCBI RefSeq database guided its taxonomic classification by calculating different genomic identity indices. All values are in percentages

PF061_03410), acetolactate synthase AlsS (*locus* PF061_04530), alpha-acetolactate decarboxylase (*locus* PF061_04535), the acetolactate synthase large subunit (*loci* PF061_07025 and PF061_09285), and the acetolactate synthase small subunit (*locus* PF061_07030).

3.2 Taxonomic assignment

We confirmed the species assignment of *Bacillus* sp. EB-40 through sequence alignment and genomics analysis. First, we aligned its genome sequence with the representative prokaryotic genomes of the NCBI Refseq database (https://ftp.ncbi.nlm.nih.gov/blast/db/; version 23/12/2023) using the BLASTn tool of BLAST version 2.13.0 [19] and considering an E-value threshold of $1e^{-10}$ to select the significant alignments. The JSpeciesWS server (https://jspecies.ribohost.com/jspeciesws/) [29] was subsequently used to compare the *Bacillus* sp. EB-40 genome with the ten best-ranked genomes among the significant alignments identified via BLASTn and calculate two average nucleotide identity (ANI) indices via BLAST (ANIb) and MUMMER (ANIm). In addition, the Genome-to-Genome Distance Calculator (GGDC) server version 3.0 (https://ggdc.dsmz.de/) [30, 31] was used to calculate the degree of digital DNA−DNA hybridization (dDDH) via BLAST alignment.

Compared with the prokaryotic representative genomes of the NCBI RefSeq database, the genome of *Bacillus* sp. EB-40 shares genomic similarity indices that surpass the thresholds recommended for species assignment, which include 95% for ANI and 70% for dDDH [32, 33], with the reference genome of *Bacillus cereus* (Table 2). These indices were even higher in comparative analyses with other *B. cereus* isolates available in GenBank, such as the genome of strain DQ01 (accession CP097351), which had values of 99.14% for ANIb, 99.30% for ANIm, and 93.80% for dDDH. Therefore, we classified *Bacillus* sp. EB-40 as a *Bacillus cereus* species.

Genome sequencing and annotation are the first steps toward complete comprehension of the genetic basis of microbial activities that directly influence the beneficial genetic and physiological responses of plants. Furthermore, transcriptomic and metabolomic studies may be conducted to elucidate how environmental stimuli influence these bacteria.

4 Conclusion

The genome *of Bacillus cereus* isolate EB-40 comprises one chromosome and three plasmids. The chromosome is a 5,613,235-bp circular double-stranded DNA with a GC content of 35.3% and 5462 genes, and the three plasmids have sizes of 215,503, 7710, and 14,472 bp with a total of 201 genes. Among the 5,386 CDSs annotated in the genome, we identified genes associated with pathways associated with its previously reported phenotypic characteristics describing it as an endophyte able to promote plant growth. Comparative genomics revealed that its genome shares similarity indices with *Bacillus cereus* genomes above the thresholds recommended for species assignment. In addition, this isolate has the capacity for effective colonization of the intercellular and intracellular spaces in the banana root system, suggesting that it is a novel isolate of *Bacillus cereus*.

Acknowledgements We thank the Núcleo de Análise de Biomoléculas (NuBioMol) of the Universidade Federal de Viçosa (UFV) for providing the facilities for data analysis. NuBioMol is financially supported by the following Brazilian agencies: Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos (Finep) and Sistema Nacional de Laboratórios e Nanotecnologias (SisNANO)/Ministério da Ciência, Tecnologia e Informação (MCTI).

Author contributions Debora Francine Gomes Silva Pereira executed the experiments. Pedro Marcus Pereira Vidigal and Samuel A. Santos executed the genome assembly and analyzed the data. Silvia Nietsche coordinated the research. Debora Francine Gomes Silva Pereira and Pedro Marcus Pereira Vidigal wrote the manuscript. Adelica A. Xavier and Marlon Cristian Toledo Pereira provided funds for the research and commented on the draft manuscript. All the authors reviewed and approved the manuscript.

Funding This research was funded by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), grant numbers CAG00390-15 and APQ00146-22, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-CAPES, Financing Code 001 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), grant numbers 427667/2016-0 and 306391/2021-0.

Availability of data and materials The datasets generated and analyzed during the present research are available from the corresponding author upon reasonable request.

Declarations

Competing interest The authors declare that they have no conflicts of interest.

Nucleotide sequence accession number The complete genome sequence of the *B. cereus* isolate EB-40 was deposited on DDBJ/EMBL/GenBank under accession number CP115717 for the chromosome and numbers CP115718, CP115719, and CP115720 for the plasmids.

References

1. Glick BR. Plant Growth-Promoting Bacteria: Mechanisms and Applications. Scientifica. 2012;2012(1):963401. https://doi.org/10.6064/2012/963401.
2. Kang SM, Radhakrishnan R, Lee KE, You YH, Ko JH, Kim JH, Lee IJ. Mechanism of plant growth promotion elicited by *Bacillus* sp. LKE15 in oriental melon. Acta Agric Scandinavica, Section B–Soil Plant Sci. 2015;65(7):637–47. https://doi.org/10.1080/09064710.2015.1040830.
3. Aloo BN, Makumba BA, Mbega ER. The potential of Bacilli rhizobacteria for sustainable crop production and environmental sustainability. Microbiol Res. 2019. https://doi.org/10.1016/j.micres.2018.10.011.
4. Kulkova I, Dobrzyński J, Kowalczyk P, Bełżecki G, Kramkowski K. Plant Growth Promotion Using *Bacillus cereus*. Int J Mol Sci. 2023. https://doi.org/10.3390/ijms24119759.
5. Hu J, Dong B, Wang D, Meng H, Li X, Zhou H. Genomic and metabolic features of *Bacillus cereus*, inhibiting the growth of *Sclerotinia sclerotiorum* by synthesizing secondary metabolites. Arch Microbiol. 2022. https://doi.org/10.1007/s00203-022-03351-5.
6. Souza SA, Xavier AA, Costa MR, Cardoso A, Pereira MC, Nietsche S. Endophytic bacterial diversity in banana'Prata Anã'(Musa spp.) roots. Genet Mol Biol. 2013;36:252–64. https://doi.org/10.1590/S1415-47572013000200016.
7. Rocha JS, Nietsche S, Pereira MC, Maria OMS, Santos RC, Xavier AA. Endophytic interaction of *Bacillus* sp. in micropropagated banana plantlets. An Acad Bras Ciênc. 2019;91(03):e20181295. https://doi.org/10.1590/0001-3765201920181295.
8. Souza GL, Silva DF, Nietsche S, Xavier AA, Pereira MC. Endophytic bacteria used as bioinoculants in micropropagated banana seedlings. Rev Bras Frutic. 2017;39(2):e-324. https://doi.org/10.1590/0100-29452017324.
9. Souza GLOD, Nietsche S, Xavier AA, Costa MR, Pereira MCT, Santos MA. Triple combinations with PGPB stimulate plant growth in micropropagated banana plantlets. Appl Soil Ecol. 2016. https://doi.org/10.1016/j.apsoil.2016.03.001.
10. De Maio N, Shaw LP, Hubbard A, George S, Sanderson ND, Swann J, Wick R, AbuOun M, Stubberfield E, Hoosdally SJ, Crook DW, Peto TEA, Sheppard AE, Bailey MJ, Read DS, Anjum MF, Walker AS, Stoesser N. Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. Microbial Genomics. 2019. https://doi.org/10.1099/mgen.0.000294.
11. Smits THM. The importance of genome sequence quality to microbial comparative genomics. BMC Genomics. 2019. https://doi.org/10.1186/s12864-019-6014-5.
12. Paterson J, Jahanshah G, Li Y, Wang Q, Mehnaz S, Gross H. The contribution of genome mining strategies to the understanding of active principles of PGPR strains. FEMS Microbiol Ecol. 2017. https://doi.org/10.1093/femsec/fiw249.
13. Krueger F, James F, Ewels P. FelixKrueger/TrimGalore: v0.6.7-DOI via Zenodo. 2021. https://doi.org/10.5281/ZENODO.5127899

14. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014. https://doi.org/10.1093/bioinformatics/btu170.
15. Salmela L, Rivals E. LoRDEC: accurate and efficient long read error correction. Bioinformatics. 2014. https://doi.org/10.1093/bioinformatics/btu538.
16. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
17. 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. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res. 2017. https://doi.org/10.1093/nar/gkx1068.
18. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Res. 2020. https://doi.org/10.1093/nar/gkaa1105.
19. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016. https://doi.org/10.1093/nar/gkw569.
20. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015. https://doi.org/10.1093/bioinformatics/btv351.
21. Aziz RK, Bartels D, Best AA, et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics. 2008. https://doi.org/10.1186/1471-2164-9-75.
22. Overbeek R. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 2005. https://doi.org/10.1093/nar/gki866.
23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Bio. 1990. https://doi.org/10.1016/S0022-2836(05)80360-2.
24. Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. Science. 1997. https://doi.org/10.1126/science.278.5338.631.
25. Moriya Y, Itoh M, Okuda S, Yohizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 2007. https://doi.org/10.1093/nar/gkm321.
26. Zeng Q, Xie J, Li Y, Gao T, Xu C, Wang Q. Comparative genomic and functional analyses of four sequenced *Bacillus cereus* genomes reveal conservation of genes relevant to plant-growth-promoting traits. Sci Rep. 2018. https://doi.org/10.1038/s41598-018-35300-y.
27. Niu DD, Liu HX, Jiang CH, Wang YP, Wang QY, Jin HL, Guo JH. The plant growth–promoting rhizobacterium *Bacillus cereus* AR156 induces systemic resistance in *Arabidopsis thaliana* by simultaneously activating salicylate-and jasmonate/ethylene-dependent signaling pathways. Mol Plant Microbe Interact. 2011. https://doi.org/10.1094/MPMI-09-10-0213.
28. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen CY, Graham M, Van Domselaar G, Stothard P. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 2023;51(W1):W484–92. https://doi.org/10.1093/nar/gkad326.
29. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci. 2009. https://doi.org/10.1073/pnas.0906412106.
30. Idriss EE, Makarewicz O, Farouk A, Rosner K, Greiner R, Bochow H, Richter T, Borris R. Extracellular phytase activity of *Bacillus amyloliquefaciens* FZB45 contributes to its plant-growth-promoting effect. Microbiology. 2022. https://doi.org/10.1099/00221287-148-7-2097.
31. Andrade LF, de Souza GLOD, Nietsche S, Xavier AA, Costa MR, Cardoso AM, Pereira MCT, Pereira DF. Analysis of the abilities of endophytic bacteria associated with banana tree roots to promote plant growth. J Microbiol. 2014. https://doi.org/10.1007/s12275-014-3019-2.
32. Duca D, Lorv J, Patten CL, Rose D, Glick BR. Indole-3-acetic acid in plant–microbe interactions. Antonie Van Leeuwenhoek. 2014. https://doi.org/10.1007/s10482-013-0095-y.
33. Carreño-Lopez R, Campos-Reales N, Elmerich C, Baca BE. Physiological evidence for differently regulated tryptophan-dependent pathways for indole-3-acetic acid synthesis in *Azospirillum brasilense*. Mol Gen Genet. 2000. https://doi.org/10.1007/s004380000340.