



The Open Microbiology Journal

Content list available at: <https://openmicrobiologyjournal.com>



RESEARCH ARTICLE

The Complete Genome Sequence of *Bacillus safensis* BRM1 Isolated from Brazilian Mangrove Sediment: A Potential Source of Biomass Converting Enzymes

Marcelo Scarduelli^{1,*}, Dieval Guizelini², Rodrigo Luis Alves Cardos³, Denny Marcel Ceccon¹, Lucélia Donatti⁴, Valter Antônio de Baura¹, Fábio de Oliveira Pedrosa¹, Luciano Fernandes Huergo¹ and Emanuel Maltempi de Souza¹

¹Programa de Pós-Graduação em Ciências – Bioquímica, Departamento de Bioquímica e Biologia Molecular, UFPR, Curitiba, PR, Brazil

²Setor de Educação Profissional e Tecnológica da UFPR, Curitiba, PR, Brazil

³Departamento de Bioquímica e Biologia Molecular, Federal University of Paraná, Brazil

⁴Departamento de Biologia Celular, UFPR, Curitiba, PR, Brazil

Abstract:

Background:

Bacillus safensis BRM1 was isolated from Brazilian mangrove sediment and selected for its ability to grow in xylan as the sole carbon source. To identify genes encoding biomass conversion enzymes, the genome of this bacterium was sequenced.

Methods:

Genome wide analysis revealed 99% nucleotide identity to the *Bacillus safensis* genome. The isolated strain was named *B. safensis* BRM1, and its genome consists of a circular chromosome of 3.74 Mb with a GC content of 41.8%. Genes encoding a plethora of hydrolytic enzymes are present in the BRM1 genome but absent from the other *B. safensis* genomes.

Results:

A total of 23 genes encoding putative cellulases or hemicellulases were identified.

Conclusion:

These data support that *B. safensis* BRM1 is an interesting candidate for the prospection of enzymes that can be applied in the conversion of cellulosic biomass to biofuel.

Keywords: *Bacillus safensis*, Xylanases, Mangrove, Genes encoding, Enzymes, Genome.

Article History	Received: July 17, 2021	Revised: November 8, 2021	Accepted: January 24, 2022
-----------------	-------------------------	---------------------------	----------------------------

1. INTRODUCTION

The genus *Bacillus* is comprised of aerobic gram-positive, rod-shaped bacteria. Species of this genus can be found in many different habitats, such as soil, water, food, diseased animals (*B. anthracis* or *B. larvae*), hot springs, and plants [1]. These bacteria have wide biotechnological applications, being the main source of a variety of hydrolytic enzymes such as amylases and proteases which are used in the industry worldwide [2]. Many *Bacillus* strains are able to produce cellulases and hemicellulases that may be applied in biofuel

production; some of these enzymes showed particular interesting properties such as a wide range of pH and temperature stability [3 - 5]. Amongst the species of the genus, *Bacillus safensis* has demonstrated the potential of producing enzymes with different biotechnology applications, including inulinases, laccases, galactosidases, oxidoreductases, and keratinases [6 - 10].

Searching for new enzymes with biotechnology potential, we collected a sample of a Brazilian mangrove, located at Paranaguá bay, Paraná (25°29'S 48°27'W). The microorganisms from the sample were isolated after growth in a selective medium, and a strain capable of growing on xylan as the sole carbon source, and in LB (Luria-Bertani) containing

* Address correspondence to this author at the Programa de Pós-Graduação em Ciências – Bioquímica, Departamento de Bioquímica e Biologia Molecular, UFPR, Curitiba, PR, Brazil; E-mail: marcelomogoda@gmail.com

5% NaCl (m/v) was isolated and named BRM1 (*BRAzilian Mangrove*). In order to identify genes encoding biomass conversion enzymes, the genome of this strain was sequenced.

2. MATERIALS AND METHODS

2.1. Sample Extraction and Strain Selection

Samples were extracted from the sediments of mangroves at Paranaguá Bay, Paraná, Brazil, during the summer season (March 14th, 2014), near Antonina city. The upper layer of sediment (around 10 cm) was discarded and around 100 grams of the surfaced material was aseptically collected and stored in ice until processing.

At the laboratory, 1 g of the sediment sample was resuspended in 5 mL of 5% NaCl. An aliquot of 200 µL of the serial dilutions at 1:25 and 1:125 was spread on an agar plate containing a selective medium for cellulases or xylanases producing microorganisms. The selective medium used had the following composition: 1.6g/L of K₂HPO₄, 0.2g/L of KH₂PO₄, 0.2 g/L of MgSO₄.7H₂O, 0.01g/L of FeSO₄.7H₂O, 0.2 g/L of yeast extract, 0.2 g/L casamino acids, 1.2x10⁻² g/L of MnSO₄.H₂O, 0.02 g/L of CaCl₂.2H₂O, 1.4x10⁻² g/L of H₃BO₃, 4.0x10⁻⁴ g/L of CuSO₄.5H₂O, 1.2x10⁻³ g/L ZnSO₄.7H₂O, 0.01 g/L of Na₂MoO₄.2H₂O, 20 g/L of NaCl, 15 g/L of agar, and 5 g/L of xylan or carboxymethyl cellulose (CMC).

After inoculation, the plates were maintained at 30°C for 4 days, and morphologically different colonies were picked, streaked onto a novel selective medium agar plate, and incubated again at 30°C for 4 days. A total of 20 isolates were selected and kept at 4 °C on solid medium or stored in skim milk at -80 °C [11]. One isolate showing higher xylan degrading activity was named BRM1 and selected for further analysis.

2.2. Enzyme Activity Measurements

To detect enzymatic activity on Petri dishes, the solid

medium was the same selective medium used during the isolate selection stage. The presence of CMC hydrolytic activity was verified by the formation of hydrolytic halo using the red congo staining [12]. For the xylanase activity, the halo could be identified directly without the need for a staining step.

For enzymatic assay in a liquid medium, the isolate was cultured aerobically in a rotary shaker at 30°C and 120 rpm. The medium composition was the same as described by COMMICHAU and collaborators [13] but diluted 2-fold in water. Different carbon sources (0.5% w/v) were added as indicated: glucose (control), cellobiose, xylan, or carboxymethyl cellulose (CMC). A 300 µL aliquot was removed daily from each cultivation flask and the cell density was monitored by measuring the optical density at 600 nm.

2.3. Genome Sequences

For genome's comparison, we used 13 complete *Bacillus* genome sequences available at public database: *B. pumilus* 145, *B. subtilis* ATCC 19217, *B. safensis* KCTC, *B. altitudinis* P-10, *B. pumilus* SAFR-032, *B. altitudinis* SGAir0031, *B. pumilus* SH-B9, *B. safensis* U14-5, *B. safensis* U17-1, *B. safensis* U41, *B. xiamenensis* VV3, *B. altitudinis* W3, and *Bacillus sp.* WP8 (Table 1).

2.4. Genome Sequencing, Annotation, and Analysis

Genomic DNA was purified using phenol-chloroform extraction of cells lysed with lysozyme and SDS. Genome sequencing of BRM1 was performed using an Illumina MiSeq Sequencer. Library construction was done with Nextera DNA Library Prep Kit, and sequencing generated 3,185,036 end-paired reads (~448 Mb), yielding a 120-fold genome coverage. The data were assembled using CLC Genomics Workbench 6.5.1 [14] and Roche GS De Novo Assembler (Newbler, release 2.8). Contigs were ordered in one scaffold using the jContigSort program [15] and the gaps were closed with the FGAP program [16] and *in house* scripts.

Table 1. Features of *Bacillus* strains used in genome comparison analysis (information taken from the NCBI GenBank platform).

Accession ^a	Organism	Isolation Source	Size (Mb)	GC (%) ^b	Identity (%) ^c
CP024204.1	<i>B. altitudinis</i> P-10	Rhizosphere of rice	3.76	41.5	87.94
CP022319.1	<i>B. altitudinis</i> SGAir0031	Air sample	3.84	41.4	87.95
CP011150.1	<i>B. altitudinis</i> W3	Raw gallnut honey	3.74	41.4	87.88
CP027116.1	<i>B. pumilus</i> 145	Sediment top	3.94	41.2	90.15
CP000813.4	<i>B. pumilus</i> SAFR-032	Spacecraft facilities	3.70	41.3	89.90
CP011007.1	<i>B. pumilus</i> SH-B9	Rhizosphere of beet	3.87	41.4	90.38
CP018197.1	<i>B. safensis</i> KCTC	Marine sponge	3.97	41.4	97.31
CP015607.1	<i>B. safensis</i> U14-5	Antarctic Lake	4.15	41.4	97.56
CP015611.1	<i>B. safensis</i> U17-1	Antarctic Lake	3.74	41.5	95.97
CP015610.1	<i>B. safensis</i> U41	Antarctic Lake	3.73	41.5	95.99
CP009749.1	<i>B. subtilis</i> ATCC 19217	Industrial strain	3.95	46.4	78.25
CP017786.1	<i>B. xiamenensis</i> VV3	Fermented rice	3.63	41.4	88.35
CP010075.1	<i>Bacillus sp.</i> WP8	Rhizosphere of wheat	3.70	41.6	95.90

^aNCBI GenBank accession number.

^bPercentage of G and C in the genome sequence, defined as 100 x (G+C)/(A+T+C+G).

^cIdentity with *B. safensis* BRM1.

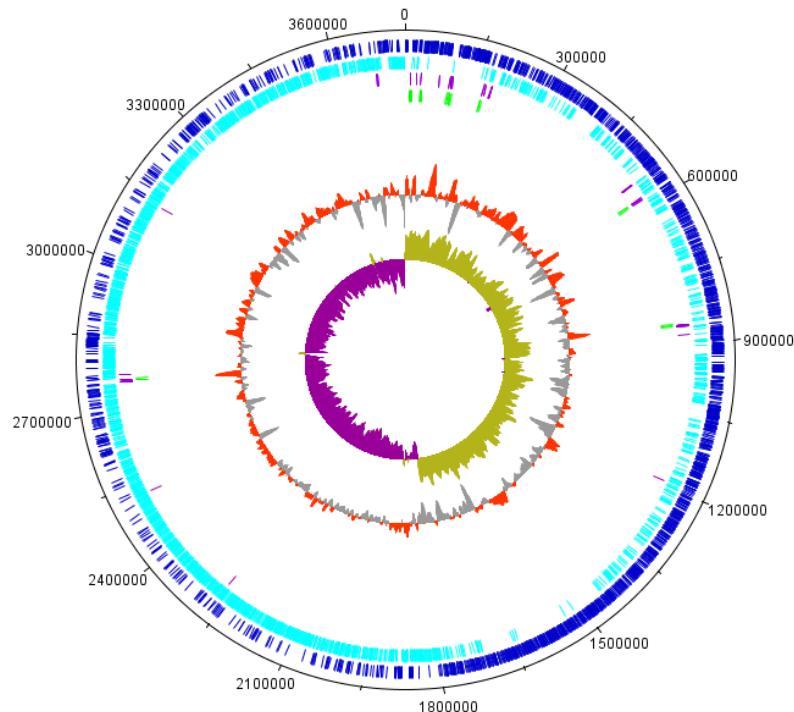


Fig. (1). Map of the *Bacillus safensis* BRM1 genome. From the outer to inner circles: ORFs on the forward strand (in dark blue) and reverse strand (light blue), tRNA genes in both strands (purple), rRNA genes in both strands (green), GC percentage plot (above average in orange, below average in grey), GC skew (lime green above average, purple below average).

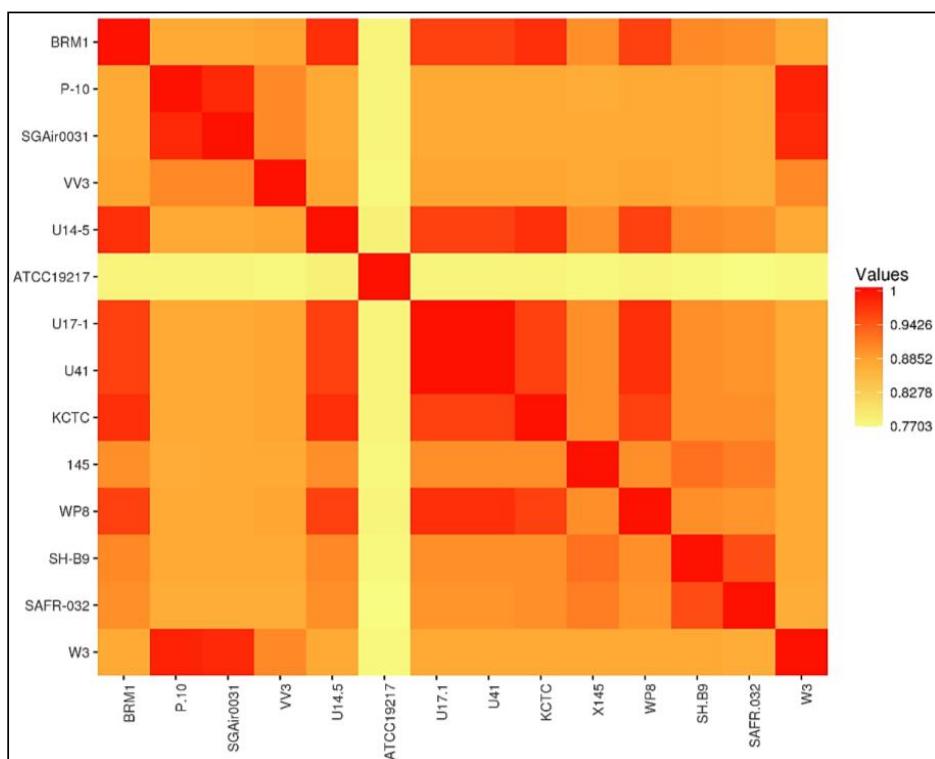


Fig. (2). Heatmap showing Average Nucleotide Identity between the 14 *Bacillus* strain genomes deposited in NCBI database. The Genbank accession numbers were as follow: *B. safensis* BRM1 (CP018100), *B. altitudinis* P-10 (CP024204.1), *B. altitudinis* SGAir0031 (CP022319.1), *B. xiamenensis* VV3 (CP017786.1), *B. safensis* U14-5 (CP015607.1), *B. subtilis* ATCC 19217 (CP009749.1), *B. safensis* U17-1 (CP015611.1), *B. safensis* U41 (CP015610.1), *B. safensis* KCTC (CP018197.1), *B. pumilus* 145 (CP027116.1), *Bacillus* sp. WP8 (CP010075.1), *B. pumilus* SH-B9 (CP011007.1), *B. pumilus* SAFR-032 (CP000813.4), and *B. altitudinis* W3 (CP011150.1).

Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline [17] on the fly during the genome deposit process at GenBank. Annotation was also performed using the RAST genome annotation service [18].

The carbohydrate-active enzymes coding genes in the strain BRM1 genome were identified using the CAZY online [19]. Ribosomal sequences were identified using RNAmmer [20] and ARAGORN [21].

The *Bacillus safensis* BRM1 genome map was generated using DNAPlotter [22]. Clusters of orthologous group (COGs) analysis were performed using the EggNog 4.5.1 software [23]. PSORTb v3.0 was used to predict the protein subcellular localization [24].

2.5. Phylogenetics Analysis

To infer phylogenetic relationships among the 14 *Bacillus* strains, we used the Average Nucleotide Identity matrix at the Kostas Lab website [25] and the results were plotted as a heatmap using the Heatmapper tool [26].

2.6. Data Availability

The genome information of *Bacillus safensis* BRM1 was deposited in NCBI under BioProject number PRJNA353942, BioSample number SAMN06030516, and GenBank accession number CP018100. The version described in this paper is the first version.

3. RESULTS AND DISCUSSION

3.1. Genome Analysis

The genome of BRM1 consists of one circular chromosome with 3,740,842 pb (Fig. 1) and a GC content of 41.80%. BRM1 strain has eight rRNA operons: eight 5S rRNAs, eight 16S rRNAs, and eight 23S rRNAs. There were 91 tRNAs, covering all 20 amino acids. The number of coding sequences (CDS) was 3,699, with an average length of 837 bp. The content of non-coding regions was 10.82%. The genome features of the BRM1 strain were similar to those of other *Bacillus safensis* strains.

3.2. Phylogenetic Analysis

Genome similarity analysis was performed using the Average Nucleotide Identity tool [25]. Table 1 shows the genome features of the 13 *Bacillus* strains included in this

study, selected due to their identity of the 16S rRNA of BRM1 and their genomes were completely sequenced. The genome sizes ranged from 3.63 Mb to 4.15 Mb, G+C content ranged from 41.2% to 46.4%, and CDS numbers ranged from 3556 to 3979. The results were plotted as a heatmap (Fig. 2). The identity between BRM1 and 13 complete *Bacillus* genomes varied from 77% to 97.56%.

The identity of BRM1 with *B. safensis* strains varied from 95.99% to 97.56%, being the highest identify degree with *B. safensis* strain U14-5 (97.56%). The next closest species (*B. pumilus*) had identities of 90% or lower. Since the identities with *B. safensis* were above 95%, these results support the classification of the BRM1 within the *B. safensis* species [25].

A genome comparison was performed at the SEED Viewer platform [27]. The results revealed a set of genes exclusively present in BRM1, encoding putative enzymes such as chitinase and ketoglutarate semialdehyde dehydrogenase, the DNA repair methyltransferase, glycerol dehydrogenase, nitrilase, and an exonuclease (Table 2).

3.3. Annotation and Protein Analysis

To annotate gene functions in *B. safensis* BRM1 genome, we use the EggNog 4.5.1 program [23], which performs a cluster of orthologous group (COGs) analysis classifying the proteins/ORFs into specific categories related to different aspects of physiology and cellular metabolism. The category distribution is depicted in the supplementary material (Table S1). Genes related to amino acid (E, 9.0%), transcription (K, 8.6%) and carbohydrate metabolism (G, 6.7%) comprised 24.3% of total protein coding regions identified. Genes for a protein belonging to unknown functions represented 24.6% of the sequences. Other categories, such as cell wall/membrane (5.3%), translation (5.1%), and ion transport and metabolism (5.1%) proteins, were also highly prevalent.

3.4. Hydrolytic Enzymes

Glycosyl hydrolases (GHs) are widely used in biomass-to-sugar conversion in food and biofuel producing processes. Analyses of BRM1 strain genome using carbohydrate-active enzyme (CAZy) database [19] identified 35 genes encoding putative biomass degradation enzymes (Table 3 and Table S2) such as cellulases, xylanases, and acetyl xylan esterases (involved in xylan degradation). The encoded enzymes were classified into 18 different families of glycosyl hydrolases (GHs) and carbohydrate esterases (CEs).

Table 2. Features of genes exclusively present the *Bacillus safensis* BRM1 (information taken from SEED Viewer and the NCBI GenBank platform).

Number ID at SEED Viewer	Probably Function at Seed Viewer	Length (pb)	Sequence ID at NCBI GenBank
1027	Chitinase	276	WP_155889681.1
893	Ketoglutarate semialdehyde dehydrogenase	927	WP_065214177.1
556	DNA repair methyltransferase	1194	WP_070326855.1
225	Glycerol dehydrogenase	1083	WP_081123921.1
2266	Nitrilase	1008	WP_170825868.1
3802	Exonuclease SbcC	1974	WP_034283567.1

Table 3. Genes for putative biomass hydrolases found in *Bacillus safensis* BRM1 genome.

Probable Activity	Conserved Domains (# of Genes Found)
<i>Cellulases</i>	-
Endo/exoglucanases	GH 5 (1), 9 (1), 48 (1), 51(1)
β -Glucosidase	GH 1 (6), 3 (1) or 42 (1)
<i>Hemicellulases</i>	
Xylanases or β -Xylosidases	GH 10 (1), 11 (1), 16 (1), 30 (1), 43 (2)
α -galactosidases	GH 4 (4)
Galactanases	GH 53 (1)
Acetyl xylan esterase	CE 4 (9), 6 (1), 7 (1), 12 (1)
-	(Total: 35 genes for putative hydrolases)

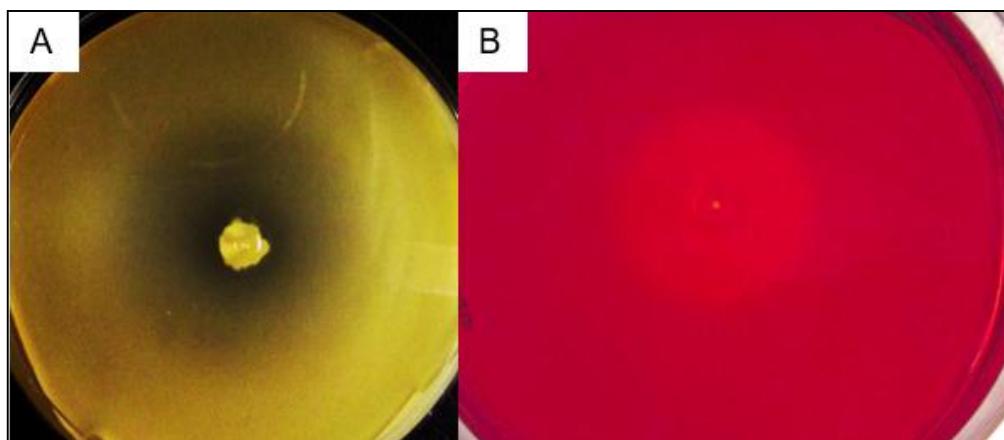


Fig. (3). Hydrolysis halo formed after 4 days at 30°C of cultivation. (A) plate containing xylan as a carbon source, indicating xylanase activity. (B) plate with CMC as a sole carbon source, showing CMCase activity. Here, the halo was revealed after the Red Congo stain.

To infer which biomass degradation enzymes could be secreted to the extracellular medium, we use the PSORTb v3.0 software (yu *et al.* 2010). As a result, 12 of those enzymes were predicted to be targeted to the extracellular medium (Table S2).

Additionally, RAST analysis [18] showed that the xylan utilization pathway is complete in the BRM1 strain genome: xylose isomerase (XylA, ARD56435.1), xylulose kinase (XylB, ARD56434.1), xylose-responsive transcription regulator (XylR, ARD56436.1), xyloside transporter (XynT, ARD56438.1), and three xylan hydrolases (putatives XynA, XynB and/or XynC: ARD56360.1, ARD56433.1 and ARD56437.1). These results show that the BRM1 strain genome encodes the enzymatic machinery to degrade recalcitrant carbon sources including xylan, cellobiose, and CMC, hence, confirming the biomass degrading potential of BRM1.

3.5. Growth On Specific Substrates

The BRM1 strain was isolated from a Brazilian mangrove sample plated on solid media containing CMC or xylan as a sole carbon source (as described in the Material and Methods section). This strain is capable of growing in 2% (w/v) NaCl using either substrate as carbon source, and the hydrolysis halo could be detected after 4 days of growth at 30°C (Fig. 3).

In order to confirm the hydrolytic potential, the BRM1 strain was cultured aerobically in a liquid medium supplemented with glucose (control), cellobiose, xylan, or carboxymethyl cellulose (CMC). For this analysis, the minimal medium CSE described by Commichau *et al.* (2008) was used, which provided better homogeneity during cell growth, without the appearance of biofilm and bacteria flocculation observed using LB or M9 medium (data not shown).

The growth curves in Fig. (4) indicate that BRM1 was capable of growing in CSE minimal medium with no addition until day 5, probably consuming the glutamate and succinate present in the broth. After this, growth stopped reaching a stationary phase. When cellobiose was added, the bacteria grew much faster, similarly to glucose, reaching maximum cell density 1-2 days of incubation. On the other hand, when xylan was used, the initial growth rate was somewhat faster than with only CSE, but the cells continued to grow even after 17 days of incubation. When CMC was used as a carbon source, the initial growth rate was marginally higher than the control condition (CSE only) and also reached maximum cell density higher than the control after 5-7 days. Since BRM1 produced a hydrolysis halo in a solid medium plate containing CMC, the growth profile in liquid medium with CMC suggests that cellulolytic activity is low, or perhaps aeration could interfere in CMCase activity.

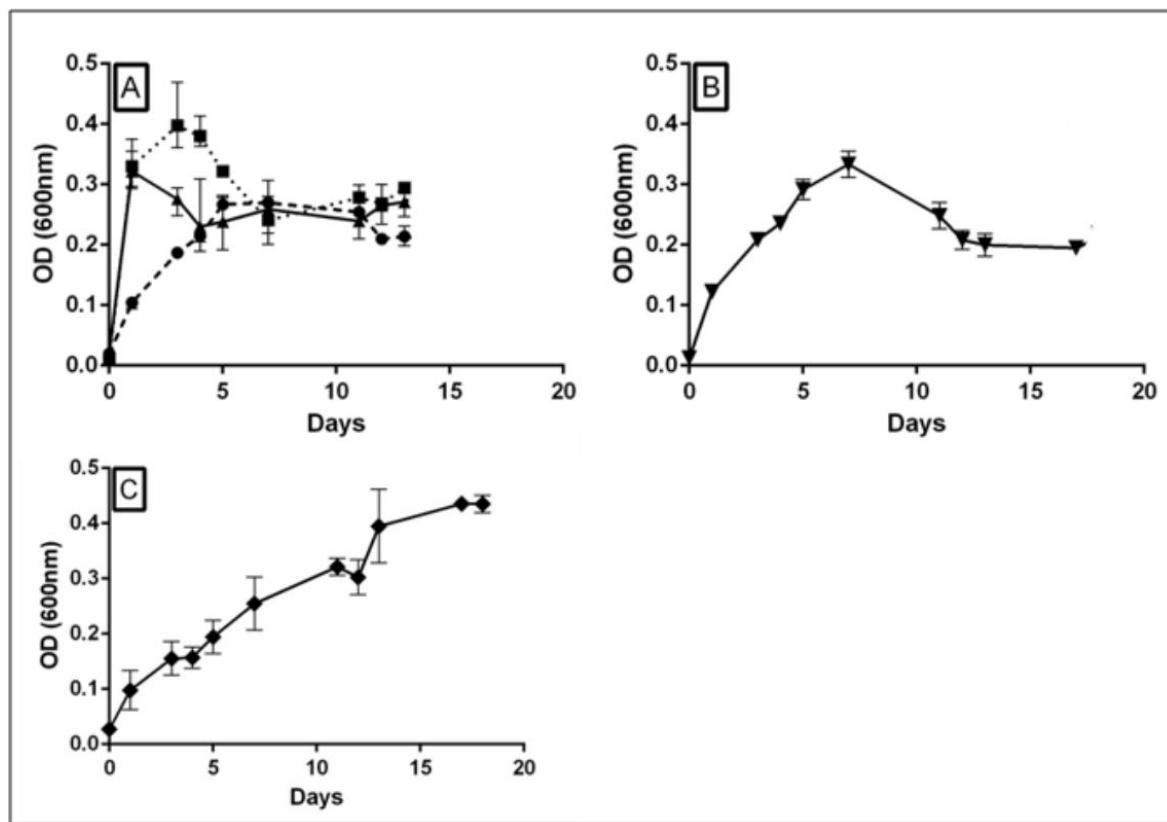


Fig. (4). *Bacillus safensis* BRM1 growth in CSE medium supplemented with cellobiose (A), CMC (B), and xylan (C), in a concentration of 0.5% (m/v). The controls were the CSE medium without carbon source (●) and with glucose at 0.5% (m/v, ■). The curve related to growth in cellobiose is marked with a triangle (▲). For growth in CMC and xylan, the points were marked with an inverted triangle (▼) and a diamond (♦), respectively. The growth was performed in triplicate ($n=3$).

CONCLUSION

Here we report the isolation of *B. safensis* BRM1 from the Brazilian mangrove soil, a strain with the capacity to grow using several recalcitrant polysaccharides. The genome sequence of this strain represents the sixth complete genome of *Bacillus safensis* deposited on the GenBank database. The other strains with complete genome described were also isolated from saline environments: Antarctic Lake and Chinese ocean.

The BRM1 strain genome encodes the enzymatic machinery to degrade different carbon sources. At least 12 enzymes potentially involved in the degradation of CMC/cellulose could be identified in the genome of BRM1; these enzymes could allow the efficient degradation of cellulose to cellobiose and then to D-glucose. Genes encoding enzymes involved in the hydrolysis of xylan to xylose were also found which would allow hemicellulose biomass degradation.

The complete genome of *Bacillus safensis* BRM1 revealed that this strain may have biotechnological potential due to the identified cellulases and hemicellulases. Future studies should be conducted to evaluate the activities of these enzymes separately and the possible usage in biomass-to-biofuel conversion processes.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

Not applicable.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

This work was supported by the National Institute of Science and Technology of Nitrogen Fixation/CNPq/MCT, Fundação Araucária and Coordination for the Improvement of Higher Education Personnel (CAPES).

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

We thank Dr. Paulo da Cunha Lana who helped during the

collecting of mangrove sediment.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

- [1] Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, Slepecky R, Hemphill E. The Prokaryotes. Bacteria: Firmicutes, Cyanobacteria. Springer: Science & Business Media 2006; 4.
- [2] Kunst F, Ogasawara N, Moszer I, et al. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 1997; 390(6657): 249-56.
[<http://dx.doi.org/10.1038/36786>] [PMID: 9384377]
- [3] Ito S. Alkaline cellulases from alkaliphilic *Bacillus*: enzymatic properties, genetics, and application to detergents. *Extremophiles* 1997; 1(2): 61-6.
[<http://dx.doi.org/10.1007/s007920050015>] [PMID: 9680303]
- [4] Khasin A, Alchanati I, Shoham Y. Purification and characterization of a thermostable xylanase from *Bacillus stearothermophilus* T-6. *Appl Environ Microbiol* 1993; 59(6): 1725-30.
[<http://dx.doi.org/10.1128/aem.59.6.1725-1730.1993>] [PMID: 8328796]
- [5] Rastogi G, Bhalla A, Adhikari A, et al. Characterization of thermostable cellulases produced by *Bacillus* and *GeoBacillus* strains. *Bioresour Technol* 2010; 101(22): 8798-806.
[<http://dx.doi.org/10.1016/j.biortech.2010.06.001>] [PMID: 20599378]
- [6] Singh RS, Singh RP, Yadav M. Molecular and biochemical characterization of a new endoinulinase producing bacterial strain of *Bacillus safensis* AS-08. *Biologia* 2013; 68: 1028-33.
[<http://dx.doi.org/10.2478/s11756-013-0259-2>]
- [7] Singh D, Sharma KK, Jacob S, Gakhar SK. Molecular docking of laccase protein from *Bacillus Safensis* DSKK5 isolated from earthworm gut: A novel method to study dye decolorization potential. *Water Air Soil Pollut* 2014; 225: 2175.
[<http://dx.doi.org/10.1007/s11270-014-2175-7>]
- [8] Nath A, Chakrabarty S, Sarkar S, Bhattacharjee C, Drioli E, Chowdhury R. Purification and characterization of β -Galactosidase synthesized from *Bacillus safensis* (JUCHE 1). *Ind Eng Chem Res* 2013; 52(33): 11663-72.
[<http://dx.doi.org/10.1021/ie4008584>]
- [9] da Fonseca FSA, Angolini CFF, Arruda MAZ, et al. Identification of oxidoreductases from the petroleum *Bacillus safensis* strain. *Biotechnol Rep (Amst)* 2015; 8: 152-9.
[<http://dx.doi.org/10.1016/j.btre.2015.09.001>] [PMID: 28352585]
- [10] Lateef A, Adelere IA, Gueguim-Kana EB, Asafa TB, Beukes LS. Green synthesis of silver nanoparticles using keratinase obtained from a strain of *Bacillus safensis* LAU 13. *Int Nano Lett* 2015; 5: 29-35.
[<http://dx.doi.org/10.1007/s40089-014-0133-4>]
- [11] Cody WL, Wilson JW, Hendrixson DR, et al. Skim milk enhances the preservation of thawed -80 degrees bacterial stocks. *J Microbiol Methods* 2008; 75(1): 135-8.
[<http://dx.doi.org/10.1016/j.mimet.2008.05.006>] [PMID: 18573555]
- [12] Carder JH. Detection and quantitation of cellulase by Congo red staining of substrates in a cup-plate diffusion assay. *Anal Biochem* 1986; 153(1): 75-9.
[[http://dx.doi.org/10.1016/0003-2697\(86\)90063-1](http://dx.doi.org/10.1016/0003-2697(86)90063-1)] [PMID: 3963384]
- [13] Commichau FM, Gunka K, Landmann JJ, Stölke J. Glutamate metabolism in *Bacillus subtilis*: gene expression and enzyme activities evolved to avoid futile cycles and to allow rapid responses to perturbations of the system. *J Bacteriol* 2008; 190(10): 3557-64.
[<http://dx.doi.org/10.1128/JB.00099-08>] [PMID: 18326565]
- [14] Gnerre S, MacCallum I, Przybylski D, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci USA* 2011; 108(4): 1513-8.
[<http://dx.doi.org/10.1073/pnas.1017351108>] [PMID: 21187386]
- [15] Guizelini D, Pedrosa FO, Tibáes JH, et al. jContigSort: A new computer application for contigs ordering. 7th International Conference of The Brazilian Association for Bioinformatics and Computational Biology: Abstract Book. Florianópolis. 2011.
- [16] Piro VC, Faoro H, Weiss VA, et al. FGAP: An automated gap closing tool. *BMC Res Notes* 2014; 7(1): 371.
[<http://dx.doi.org/10.1186/1756-0500-7-371>] [PMID: 24938749]
- [17] Angiuoli SV, Gussman A, Klimke W, et al. Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. *OMICS* 2008; 12(2): 137-41.
[<http://dx.doi.org/10.1089/omi.2008.0017>] [PMID: 18416670]
- [18] Aziz RK, Bartels D, Best AA, et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* 2008; 9(1): 75.
[<http://dx.doi.org/10.1186/1471-2164-9-75>] [PMID: 18261238]
- [19] Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res* 2009; 37(Database issue): D233-8.
[<http://dx.doi.org/10.1093/nar/gkn663>] [PMID: 18838391]
- [20] Lagesen K, Hallin P, Rodland EA, Staerfeldt H-H, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007; 35(9): 3100-8.
[<http://dx.doi.org/10.1093/nar/gkm160>] [PMID: 17452365]
- [21] Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 2004; 32(1): 11-6.
[<http://dx.doi.org/10.1093/nar/gkh152>] [PMID: 14704338]
- [22] Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics* 2009; 25(1): 119-20.
[<http://dx.doi.org/10.1093/bioinformatics/btn578>] [PMID: 18990721]
- [23] Huerta-Cepas J, Forslund K, Coelho LP, et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol Biol Evol* 2017; 34(8): 2115-22.
[<http://dx.doi.org/10.1093/molbev/msx148>] [PMID: 28460117]
- [24] Yu NY, Wagner JR, Laird MR, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 2010; 26(13): 1608-15.
[<http://dx.doi.org/10.1093/bioinformatics/btq249>] [PMID: 20472543]
- [25] Rodriguez-R LM, Konstantinidis KT. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes 2016.
[<http://dx.doi.org/10.7287/peerj.preprints.1900v1>]
- [26] Babicki S, Arndt D, Marcu A, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* 2016; 44(W1): W147-53.
[<http://dx.doi.org/10.1093/nar/gkw419>] [PMID: 27190236]
- [27] Overbeek R, Begley T, Butler RM, et al. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 2005; 33(17): 5691-702.
[<http://dx.doi.org/10.1093/nar/gki866>] [PMID: 16214803]

© 2022 Scarduell *et al.*

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.