

High-Quality Draft Genome Sequence of *Bacillus amyloliquefaciens* Strain 629, an Endophyte from *Theobroma cacao*

Brena M. M. SantAnna,^a Phellippe P. A. Marbach,^b Marcelo Rojas-Herrera,^c Jorge T. De Souza,^d Milton R. A. Roque,^a Artur T. L. Queiroz^e

Universidade Federal da Bahia (UFBA), Salvador, Brazil^a; Universidade Federal do Recôncavo da Bahia (UFRB), Cruz das Almas, Brazil^b; Centro de Genômica y Bioinformática, Universidad Mayor, Santiago, Chile^c; Universidade Federal de Lavras (UFLA), Lavras, Brazil^d; Centro de Pesquisas Gonçalo Moniz (CPqGM)-FIOCRUZ, Salvador, Brazil^e

***Bacillus amyloliquefaciens* strain 629 is an endophyte isolated from *Theobroma cacao* L. Here, we report the draft genome sequence (3.9 Mb) of *B. amyloliquefaciens* strain 629 containing 16 contigs (3,903,367 bp), 3,912 coding sequences, and an average 46.5% G+C content.**

Received 29 September 2015 Accepted 1 October 2015 Published 19 November 2015

Citation SantAnna BMM, Marbach PPA, Rojas-Herrera M, De Souza JT, Roque MRA, Queiroz ATL. 2015. High-quality draft genome sequence of *Bacillus amyloliquefaciens* strain 629, an endophyte from *Theobroma cacao*. *Genome Announc* 3(6):e01325-15. doi:10.1128/genomeA.01325-15.

Copyright © 2015 SantAnna et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Artur T. L. Queiroz, artur.queiroz@bahia.fiocruz.br.

Bacilli are frequently isolated as endophytes and are common components of the microbiota of several plant species (1, 2). Strain 629 was isolated from a healthy *Theobroma cacao* tree and was initially identified as *Bacillus subtilis* (3), but further analysis based on *gyrB* and *recA* sequences revealed that its true identity is *Bacillus amyloliquefaciens* (4). This isolate colonizes different host and plant tissues under both sterile and nonsterile conditions and promotes plant growth (3, 4). Strain 629 produces the lipopeptides iturin, fengicin, and surfactin and volatile organic compounds that may be active in the biocontrol of several fungal plant pathogens (unpublished data) and pathogenic bacteria, including *Curtobacterium flaccumfaciens* pv. *flaccumfaciens* (5). Furthermore, *B. amyloliquefaciens* 629 is currently being used as a model to study endophytic colonization (4). This strain is deposited in the Biological Institute Culture Collection of Phytopathogenic Bacteria (IBSBF) (Campinas, São Paulo, Brazil) under accession no. IBSBF-3106. This collection is registered with the World Data Centre for Microorganisms collection under no. WDCM-110.

Genomic DNA from isolate 629 was extracted and sequenced using the Ion Torrent PGM platform (Life Technologies) 318 chip. A total of 7,567,586 reads with an average length of 330 pb were obtained. All reads were assembled to an initial draft genome of 3,866,991 nucleotides at 443-fold coverage using the SPAdes Genome Assembler version 3.5.0, generating 129 unoriented contigs, with a G+C content of 46.5%, (N_{50} : 285,363 bp).

Contigs were ordered using CONTIGuator 2.3 (<http://contiguator.sourceforge.net/>) (6) with the *B. Amyloliquefaciens* CC178 genome, the closest available, as a reference (GenBank accession no. CP006845.1). Subsequently, 34 contigs with 3.8 Mb were aligned with the reference genome to order the contigs. A total of 95 contigs (only 9 > 600 bp) corresponding to 29,876 nucleotides were not mapped to the reference genome. These sequences were identified as redundant contigs, according to BLAST results, and were removed from the assembly. To solve the repetitive sequences and the remaining gaps the MapRepeat pipeline

(7) was used, resulting in the final high-quality draft genome sequence with 16 contigs, containing 3,903,367 bp.

Genome annotation was performed with RAST version 2.0 server (8). The genome of strain 629 is composed of 4,013 predicted genes, including 3,912 protein-coding sequences, 82 tRNAs, and 19 copies of the genes for 5S, 16S, and 23S rRNA. The genome of strain 629 is closely related to that of *B. amyloliquefaciens* CC178 with an identity of 99% (97% coverage) and also has a similar numbers of predicted genes (9).

Subsequent analysis of the genome content of *B. amyloliquefaciens* 629 and its comparison with phylogenetically related strains will help to determine key aspects of its interaction with the environment, plants, and other microorganisms.

Nucleotide sequence accession numbers. The *Bacillus amyloliquefaciens* strain 629 whole-genome shotgun (WGS) project has been deposited at DDBJ/EMBL/GenBank under the accession no. **LGYP00000000**. The version described in this paper is the first version, LGYP01000000, and consists of sequences LGYP01000001 to LGYP01000016.

ACKNOWLEDGMENTS

We thank CAPES and CNPq for financial support through the PVE project 61/2011, Zayda M. Moreira, Artur Silva, Rommel T.J. Ramos, and Luis Pacheco for DNA extraction, sequencing, and assembly assistance. J.T.S. acknowledges his productivity scholarship.

REFERENCES

- White JF, Torres MS, Sullivan RF, Jabbour RE, Chen Q, Tadych M, Irizarry I, Bergen MS, Havkin-Frenkel D, Belanger FC. 2014. Occurrence of *Bacillus amyloliquefaciens* as a systemic endophyte of vanilla orchids. *Microsc Res Tech* 77:874–885. <http://dx.doi.org/10.1002/jemt.22410>.
- Wang X, Liang G. 2014. Control efficacy of an endophytic *Bacillus amyloliquefaciens* strain BZ6-1 against peanut bacterial Wilt, *Ralstonia solanacearum*. *BioMed Res Int* 2014:465435. <http://dx.doi.org/10.1155/2014/465435>.
- Leite HAC, Silva AB, Gomes FP, Gramacho KP, Faria JC, De Souza JT, Loguercio LL. 2013. *Bacillus subtilis* and *Enterobacter cloacae* endophytes from healthy *Theobroma cacao* L. trees can systemically colonize seedlings

- and promote growth. *Appl Microbiol Biotechnol* 97:2639–2651. <http://dx.doi.org/10.1007/s00253-012-4574-2>.
4. Moreira ZM, Duarte EAA, Oliveira TAS, Monteiro FP, Loguercio LL, De Souza JT. 2015. Host and tissue preferences of *Enterobacter cloacae* and *Bacillus amyloliquefaciens* for endophytic colonization. *Afr J Microbiol Res* 9:1352–1356. <http://dx.doi.org/10.5897/AJMR2015.7475>.
 5. Martins SJ, de Medeiros FHV, de Souza RM, de Resende MLV, Ribeiro PM, Jr. 2013. Biological control of bacterial wilt of common bean by plant growth-promoting rhizobacteria. *Biol Contr* 66:65–71. <http://dx.doi.org/10.1016/j.biocontrol.2013.03.009>.
 6. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A. 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med* 6:11. <http://dx.doi.org/10.1186/1751-0473-6-11>.
 7. Mariano DC, Pereira FL, Ghosh P, Barh D, Figueiredo HC, Silva A, Ramos RT, Azevedo VA. 2015. MapRepeat: an approach for effective assembly of repetitive regions in prokaryotic genomes. *Bioinformatics* 11: 276–279. <http://dx.doi.org/10.6026/97320630011276>.
 8. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2013. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucl Acid Res* 42: D206–D214. <http://dx.doi.org/10.1093/nar/gkt1226>.
 9. Kim BY, Lee SY, Ahn JH, Song J, Kim WG, Weon HY. 2015. Complete genome sequence of *Bacillus amyloliquefaciens* subsp. *plantarum* CC178, a phyllosphere bacterium antagonistic to plant pathogenic fungi. *Genome Announc* 3(1):e01368-14. <http://dx.doi.org/10.1128/genomeA.01368-14>.