Draft Genome Sequence of *Pediococcus acidilactici* Strain LPBC161, Isolated from Mature Coffee Cherries during Natural Fermentation

Elisângela S. M. Muynarsk,[a] Gilberto V. de Melo Pereira,[a] Dany Mesa,[b] Vanete Thomaz-Soccol,[a] Júlio C. Carvalho,[a] Maria Giovana B. Pagnoncelli,[a] Carlos Ricardo Soccol[a]

[a]Department of Bioprocess Engineering and Biotechnology, Federal University of Paraná, Curitiba, Paraná, Brazil
[b]Department of Biochemistry, Federal University of Paraná, Curitiba, Paraná, Brazil

ABSTRACT Here, we report the draft genome sequence of *Pediococcus acidilactici* strain LPBC161, a lactic acid bacterium isolated from mature coffee cherries in Brazil. The genome sequence of *P. acidilactici* LPBC161 provides valuable information on the mechanisms of adaptation and metabolism of lactic acid bacteria (LAB) in the environment and stressor factors of coffee processing.

Mature coffee cherries contain nutrients that favor the growth of lactic acid bacteria (LAB), which rapidly proliferate and initiate desirable lactic fermentation. These lactic acid-producing bacteria contribute to the demucilage process of coffee pulp, which is necessary for storage and transport of the coffee beans (1). The draft genome sequence of a LAB, isolated from coffee fruits in Brazil, was sequenced and analyzed. The isolate was obtained by plating depulped coffee fruits, collected during natural fermentation, onto MRS agar (2). Genomic DNA was extracted using the phenol-chloroform method (3) and was used in a PCR for amplification of the 16S rRNA gene according to the method described by Cruz et al. (4). The isolate was identified as *Pediococcus acidilactici* and named LPBC161.

The genomic DNA was quantified by Qubit (Invitrogen) fluorimetry, diluted, and used for the construction of genomic DNA sequencing libraries using a Nextera XT DNA sample preparation kit (Illumina), according to the manufacturer's instructions. Final library quality control was performed using the 2100 bioanalyzer (Agilent Technologies). The libraries were diluted to 500 pM and pooled. This pool was quantified by quantitative PCR (qPCR) using a Kapa Biosystems preparation kit, and a 17.5 pM pooled library was sequenced with the Illumina MiSeq 500 v2 platform (paired-end, 250-bp reads). Overall, 5,795,728 reads were obtained, giving a 32-fold coverage for the strain sequenced. The Illumina reads were subjected to adapter trimming and quality filtering by Trimmomatic (5) with the following parameters: read value, <Q20; and read length, 150 bp. The quality of trimmed reads was checked by FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). The sequence data set was *de novo* assembled using the SPAdes program version 3.12 (6) with the careful mode enabled. Afterward, the genome sequences were submitted to the Rapid Annotations using Subsystems Technology (RAST) server annotation pipeline (7) to identify putative coding sequences (CDS) and provide an initial automatic annotation. When necessary, annotations were manually curated using Artemis (Sanger Institute, Cambridge, UK).

The draft genome of *P. acidilactici* LPBC161 was 1,960,506 bp, with a G+C content of 42.2% and N_{50} value of 18,140 bp. The genome annotation identified 2,019 coding sequences (CDS). Based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation, the core genome consists, in large part, of CDS for enzymes, transporters, and transcription factors. The LPBC161 strain can be classified as a facultative heterofer-

mentative bacterium, possessing all enzymes required for glycolysis, phosphoketolase, and pentose-phosphate pathways. A genome comparison by BLASTn showed high identity with *P. acidilactici* SRCM100424 and *P. acidilactici* SRCM101189. Seventy-eight unique genes, including genes related with the phosphoenolpyruvate (PEP)-dependent phospho-transferase system and transcription regulation, were reported. *P. acidilactici* LPBC161 features a high number of specific genes encoding the PEP-phosphotransferase system for sorbose, mannose, and fructose, which are present in coffee pulp (8). In addition, coding sequences for enzymes involved in the use of various sugars are present, such as trehalose-6-phosphate hydrolase, alpha-L-rhamnosidase, 2-hydroxy-3-oxopropionate reductase, pyruvate kinase, glycerate kinase, beta-glucosidase, 6-phospho-beta-glucosidase, glucose-6-phosphate 1-dehydrogenase, 6-phosphofructokinase, galactosamine-6-phosphate isomerase AgaS, L-ribulose-5-phosphate 4-epimerase, and L-arabinose isomerase. *P. acidilactici* LPBC161 contains a range of genes encoding stress-related proteins, such as catalase, thiol peroxidase, glutathione reductases, NADH-oxidases, NADH peroxidises, F_oF_1-ATPase, and alkaline shock proteins. These data open new horizons for further research into the mechanisms of adaptation and metabolism of LAB and their influence on coffee processing and quality.

Data availability. The complete genome sequence for *P. acidilactici* LPBC161 has been deposited in DDBJ/ENA/GenBank under the accession number SAXR00000000. Raw sequencing data have been deposited in the SRA database under accession number SRX5328662. The genome sequence was also deposited at the National Center for Biotechnology Information (NCBI) under the BioSample accession number SAMN10690227 and assembly number ASM402229.

REFERENCES

1. de Carvalho Neto DP, de Melo Pereira GV, Finco AMO, Letti LAJ, da Silva BJG, Vandenberghe LPS, Soccol CR. 2018. Efficient coffee beans mucilage layer removal using lactic acid fermentation in a stirred-tank bioreactor: kinetic, metabolic and sensorial studies. Food Biosci 26:80–87. https://doi.org/10.1016/j.fbio.2018.10.005.

2. de Melo Pereira GV, de Carvalho Neto DP, Medeiros ABP, Soccol VT, Neto E, Woiciechowski AL, Soccol CR. 2016. Potential of lactic acid bacteria to improve the fermentation and quality of coffee during on-farm processing. Int J Food Sci Technol 51:1689–1695. https://doi.org/10.1111/ijfs.13142.

3. de Carvalho Neto DP, de Melo Pereira GV, de Carvalho JC, Thomaz Soccol V, Soccol CR. 2017. High-throughput rRNA gene sequencing reveals high and complex bacterial diversity associated with Brazilian coffee bean fermentation. Food Technol Biotechnol 56:90–95. https://doi.org/10.17113/ftb.56.01.18.5441.

4. Cruz LM, de Souza EM, Weber OB, Baldani JI, Döbereiner J, Pedrosa FO. 2001. 16S ribosomal DNA characterization of nitrogen-fixing bacteria isolated from banana (*Musa* spp.) and pineapple (*Ananas comosus* (L.) Merril). Appl Environ Microbiol 67:2375–2379. https://doi.org/10.1128/AEM.67.5.2375-2379.2001.

5. 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.

6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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.

7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

8. Pereira GVM, de Carvalho Neto DP, Magalhães Júnior AI, Vásquez ZS, Medeiros ABP, Vandenberghe LPS, Soccol CR. 2019. Exploring the impacts of postharvest processing on the aroma formation of coffee beans—a review. Food Chem 272:441–452. https://doi.org/10.1016/j.foodchem.2018.08.061.