



Draft Genome Sequences of *Pseudomonas* sp. Strain 382 and *Pantoea coffeiphila* 342, Endophytic Bacteria Isolated from Brazilian Guarana [*Paullinia cupana* (Mart.) Ducke]

Kátia Aparecida de Siqueira,^a Rhavena Graziela Liotti,^c Tiago Antônio de Oliveira Mendes,^b Marcos Antônio Soares^a

^aDepartment of Botany and Ecology, Laboratory of Biotechnology and Microbial Ecology, Institute of Biosciences, Federal University of Mato Grosso, Cuiabá, Mato Grosso, Brazil

^bDepartment of Biochemistry and Molecular Biology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil

^cFederal Institute of Mato Grosso, Cáceres, Mato Grosso, Brazil

ABSTRACT *Pseudomonas* sp. strain 382 and *Pantoea coffeiphila* 342 are two endophytic bacterial strains isolated from *Paullinia cupana* (guarana) seeds. Their draft genome sizes were 5.96 and 6.38 Mbp, with 315 and 266 scaffolds and 52% and 62% GC content, respectively.

We have selected two endophytes, *Pseudomonas* sp. strain 382 and *Pantoea coffeiphila* 342, of the seeds of *Paullinia cupana* (1), an Amazonian plant of economic interest.

First, the DNA was extracted using a PowerSoil DNA isolation kit, and the libraries were prepared using a Nextera DNA library preparation kit. After quantification by quantitative PCR (qPCR) and size quality control of libraries using the Agilent 1200 Bioanalyzer (Agilent, USA), the DNA samples were sequenced using the HiSeq 4000 sequencer (Illumina, USA) in the 150-bp paired-end mode. Adapters and low-quality sequences were removed using the software Trimmomatic (2).

The whole-genome sequences of *P. coffeiphila* 342 and *Pseudomonas* sp. 382 were *de novo* assembled using SPAdes (3), resulting in 315 and 66 scaffolds, respectively. Their genomes were represented in single linear chromosomes with, respectively, 5,967,657 and 6,385,645 bp, 52% and 62% GC content, and N_{50} values of 288,204 bp (mean read length, 555,521 bp) and 178,683 bp (mean read length, 585,472 bp). Genome completeness analysis with BUSCO (4) resulted in 99.5% complete (C) (0.4% duplicated [D]), 0.0% fragmented (F), 0.5% missed (M), and 452 genes (n) for *P. coffeiphila* 342; and 98% C (2.0% D), 0.2% F, 1.8% M, and 1,452 genes (n) for *Pseudomonas* sp. 382, indicating that the assembled genome sequences covered most of the coding regions.

Annotation of the genome sequence using the Rapid Annotations using Subsystem Technology (RAST) server (5), followed by tRNA and rRNA detection using ARAGORN (6) and RNAmmer (7), respectively, identified 5,528 coding sequences, 79 tRNAs, and 9 rRNAs in *P. coffeiphila* 342 and 5,998 coding sequences, 77 tRNAs, and 11 rRNAs in *Pseudomonas* sp. 382.

Accession number(s). The complete genome sequences of *Pantoea coffeiphila* 342 and *Pseudomonas* sp. strain 382 were deposited in GenBank under the accession numbers [PDET00000000](#) and [PEMX00000000](#), respectively. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This study was conducted using resources from different funding agencies, including FAPEMAT, CNPq, CAPES, and the National Institute of Science and Technology in Wetlands (INAU).

Received 6 March 2018 **Accepted** 17 March 2018 **Published** 19 April 2018

Citation de Siqueira KA, Liotti RG, Mendes TADO, Soares MA. 2018. Draft genome sequences of *Pseudomonas* sp. strain 382 and *Pantoea coffeiphila* 342, endophytic bacteria isolated from Brazilian guarana [*Paullinia cupana* (Mart.) Ducke]. *Genome Announc* 6: e00287-18. <https://doi.org/10.1128/genomeA.00287-18>.

Copyright © 2018 de Siqueira et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Kátia Aparecida de Siqueira, katisiqueira.kas@gmail.com, or Marcos Antônio Soares, drmasoares@gmail.com.

This research was conducted with the help of the Cluster UFV (Universidade Federal de Viçosa).

REFERENCES

1. Liotti RG, da Silva Figueiredo MI, da Silva GF, de Mendonça EAF, Soares MA. 2018. Diversity of cultivable bacterial endophytes in *Paullinia cupana* and their potential for plant growth promotion and phytopathogen control. *Microbiol Res* 207:8–18. <https://doi.org/10.1016/j.micres.2017.10.011>.
2. 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>.
3. 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>.
4. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.
5. 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>.
6. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <https://doi.org/10.1093/nar/gkm160>.
7. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. <https://doi.org/10.1093/nar/gkh152>.