CrossMark
click for updates

Draft Genome Sequence of *Burkholderia gladioli* Strain UCD-UG_CHAPALOTE (Phylum *Proteobacteria*)

Cassandra L. Ettinger,[a] Hanan R. Shehata,[b,c] David Johnston-Monje,[b*] Manish N. Raizada,[b] Jonathan A. Eisen[a,d,e]

Genome Center, University of California Davis, Davis, California, USA[a]; Department of Plant Agriculture, University of Guelph, Guelph, Ontario, Canada[b]; Department of Microbiology, School of Pharmacy, Mansoura University, Mansoura, Egypt[c]; Department of Evolution and Ecology, University of California Davis, Davis, California, USA[d]; Department of Medical Microbiology and Immunology, University of California Davis, Davis, California, USA[e]

* Present address: David Johnston-Monje, Symbiota, Cambridge, Massachusetts, USA.

Here, we present the draft genome of *Burkholderia gladioli* strain UCD-UG_CHAPALOTE. This strain is an endophyte isolated from surface sterilized seeds of an ancient Mexican landrace of corn, Chapalote. The genome contains 8,527,129 bp in 109 scaffolds.

Address correspondence to Jonathan A. Eisen, jaeisen@ucdavis.edu.

Members of *Burkholderia gladioli* (formerly *Pseudomonas marginata*) are aerobic Gram-negative rod-shaped soil bacteria (1). Representatives of this species include plant pathogens, opportunistic human pathogens, and endophytes (2–4). *B. gladioli* strain UCD-UG_CHAPALOTE is an endophyte that was isolated from surface sterilized seeds of an ancient landrace of Mexican corn (*Zea mays* subsp. *mays*) known as Chapalote originating from the State of Sinaloa (5). The endophyte was isolated at the University of Guelph, Guelph, Ontario, Canada during August 2008 as part of an ongoing project investigating the antifungal properties of endophytic bacteria (6). Genomic DNA was extracted at the University of Guelph using a bacterial genomic DNA isolation kit (catalog no. 17900; Norgen Biotek Corp) according to the manufacturer's recommendations and DNA was ethanol precipitated before being shipped to the University of California Davis for library preparation, sequencing, and analysis. Illumina paired-end libraries were made using a Nextera DNA sample preparation kit (Illumina) and libraries were sequenced on an Illumina MiSeq with a read length of 250 bp.

A total of 3,858,496 paired-end reads were produced, and after quality trimming and error correction, 3,729,095 high-quality reads were retained. Sequence processing and assembly were performed using the A5 assembly pipeline (version A5-miseq 20140604) following the workflow described by Dunitz et al. (7, 8). The A5 assembly pipeline automates the processes of data cleaning, error correction, contig assembly, scaffolding, and quality control. The assembly resulted in 150 contigs that were contained in 109 scaffolds (minimum, 508 bp; maximum, 471,925 bp; N_{50}, 148,354 bp). The final assembly contained 8,527,129 bp with a GC content of 67.75% and a median coverage of 98×. Genome completeness was assessed using Phylosift software (version 1.0.1), which searches for 37 highly conserved, single copy marker genes (9, 10). All of the marker genes were found in the final assembly.

Automated annotation was performed using the RAST server (11). *B. gladioli* strain UCD-UG_CHAPALOTE contains 7,721 predicted protein coding sequences and 74 predicted non-coding RNAs. Previous sequencing of the 16S rRNA gene identified this isolate as a member of *B. gladioli*. A full-length (1,490-bp) 16S rRNA gene sequence was obtained from the RAST annotation and was used to verify the identity of the *Burkholderia* species by aligning it to 16S rRNA gene sequences from 59 *Burkholderia* isolates and an archaea outgroup from the Ribosomal Database Project (RDP) (12). This alignment was used to construct a phylogenetic tree using FastTree 2 (http://dx.doi.org/10.6084/m9.figshare.1245105) (13). Additionally, the full-length 16S sequence was found to be 100% identical to publically available *B. gladioli* 16S sequences confirming the species identity of this isolate.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. JRGO00000000. The version described in this paper is version JRGO01000000. The raw Illumina reads are available at ENA/SRA accession no. PRJEB7719 (ERP008658).

REFERENCES

1. **Yabuuchi E, Kosako Y, Oyaizu H, Yano I, Hotta H, Hashimoto Y, Ezaki T, Arakawa M.** 1992. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new

genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. Microbiol Immunol **36:**1251–1275. http://dx.doi.org/10.1111/j.1348-0421.1992.tb02129.x.

2. **Estrada P, Mavingui P, Cournoyer B, Fontaine F, Caballero-mellado J.** 2002. A N2-fixing endophytic *Burkholderia* sp. associated with maize plants cultivated in Mexico. Can J Microbiol **294:**285–294.

3. **Mcinroy JA, Kloepper JW.** 1995. Survey of indigenous bacterial endophytes from cotton and sweet corn. Plant Soil **173:**337–342. http://dx.doi.org/10.1007/BF00011472.

4. **Coenye T, Vandamme P.** 2003. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. Environ Microbiol **5:**719–729. http://dx.doi.org/10.1046/j.1462-2920.2003.00471.x.

5. **Turrent A, Serratos J.** 2004. Context and background on maize and its wild relatives in Mexico. *In* Sarukhán J, Raven P (ed), Maize and biodiversity: the effects of transgenic maize in Mexico. Secretariat of the Commission for Environmental Cooperation of North America, Montreal, Quebec, Canada.

6. **Johnston-Monje D, Raizada MN.** 2011. Conservation and diversity of seed associated endophytes in *Zea* across boundaries of evolution, ethnography and ecology. PLoS One **6:**e20396. http://dx.doi.org/10.1371/journal.pone.0020396.

7. **Tritt A, Eisen JA, Facciotti MT, Darling AE.** 2012. An integrated pipeline for *de novo* assembly of microbial genomes. PLoS One **7:**e42304. http://dx.doi.org/10.1371/journal.pone.0042304.

8. **Dunitz MI, Lang JM, Jospin G, Darling AE, Eisen JA, Coil DA.** 2014. Swabs to genomes: a comprehensive workflow. Peer J **2:**e453v1.

9. **Darling AE, Jospin G, Lowe E, Matsen FA, Bik HM, Eisen JA.** 2014. PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ **2:**e243. http://dx.doi.org/10.7717/peerj.243.

10. **Wu D, Jospin G, Eisen JA.** 2013. Systematic identification of gene families for use as "markers" for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their major subgroups. PLoS One **8:**e77033. http://dx.doi.org/10.1371/journal.pone.0077033.

11. **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. http://dx.doi.org/10.1186/1471-2164-9-75.

12. **Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM.** 2014. Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res **42:**D633–D642. http://dx.doi.org/10.1093/nar/gkt1244.

13. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS One **5:**e9490. http://dx.doi.org/10.1371/journal.pone.0009490.