



**HAL**  
open science

## Complete genome of *Bradyrhizobium* sp. strain BDV5419, representative of Australian genospecies L

Christine Oger-Desfeux, Jerome Briolay, P. Oger, Bénédicte Lafay

### ► To cite this version:

Christine Oger-Desfeux, Jerome Briolay, P. Oger, Bénédicte Lafay. Complete genome of *Bradyrhizobium* sp. strain BDV5419, representative of Australian genospecies L. *Microbiology Resource Announcements*, 2021, 10 (8), pp.e00042-21. 10.1128/MRA.00042-21 . hal-03138171

**HAL Id: hal-03138171**

**<https://hal.science/hal-03138171>**


Submitted on 26 Feb 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Complete Genome Sequence of *Bradyrhizobium* sp. Strain BDV5419, Representative of Australian Genospecies L

Christine Oger-Desfeux,<sup>a</sup> Jérôme Briolay,<sup>b</sup> Philippe M. Oger,<sup>c</sup>  Bénédicte Lafay<sup>d,e</sup>

<sup>a</sup>Université de Lyon, Université Claude Bernard-Lyon 1, PRABI-ASMB, FR3728 Bio-Environnement et Santé, Villeurbanne, France

<sup>b</sup>Université de Lyon, Université Claude Bernard-Lyon 1, CNRS, DTAMB, FR3728 Bio-Environnement et Santé, Villeurbanne, France

<sup>c</sup>Université de Lyon, INSA de Lyon, CNRS, UMR5240, Villeurbanne, France

<sup>d</sup>Université de Lyon, École Centrale de Lyon, CNRS, UMR5005, Laboratoire Ampère, Écully, France

<sup>e</sup>Université de Lyon, Université Claude Bernard-Lyon 1, CNRS, UMR5558, Laboratoire de Biométrie et Biologie Évolutive, Villeurbanne, France

**ABSTRACT** We report the complete genome sequence of *Bradyrhizobium* sp. strain BDV5419, representative of *Bradyrhizobium* genospecies L, which symbiotically associates with the Australian native legume *Hardenbergia violaceae* and is expected to represent a novel *Bradyrhizobium* species. The complete genome sequence provides a genetic reference for this Australian genospecies.

**B** *Bradyrhizobium* strain BDV5419 was isolated in 1995 from a root nodule of *Hardenbergia violaceae* (Fabaceae, Faboideae, Phaseoleae) collected on Black Mountain in Canberra, Australian Capital Territory, Australia (−35°16′1.20″S, 149°05′60.00″E), in the course of a survey of rhizobia associated with native shrubby legumes in southeastern Australia (1). It is the representative strain of *Bradyrhizobium* genospecies L, which may constitute a novel *Bradyrhizobium* species related to *Bradyrhizobium elkanii*, based on small-subunit (SSU) rRNA gene sequence analysis and phylogeny (1). We sequenced its genome to further the characterization of this genospecies and to gain insight regarding its phylogenetic position and taxonomy within the *Bradyrhizobium* genus.

Strain BDV5419 was grown from a lyophilized stock in 30 ml of yeast extract-mannitol broth (2) at 25°C and 200 rpm for 5 days. Genomic DNA was prepared, as described, by successive phenol-chloroform extractions (3). DNA quantification and quality control were performed using a NanoDrop spectrophotometer, a Qubit 4 fluorometer, and agarose gel electrophoresis. The same DNA was used for Nanopore and Illumina sequencing. Illumina libraries were obtained using the Nextera XT kit following the manufacturer's instructions, starting with 1 ng of genomic DNA, and were analyzed by paired-end (2 × 300-bp) sequencing on a MiSeq instrument. Poor-quality regions (Q < 30) of raw reads were removed using Sickle v1.33 (4). Long reads were obtained with an Oxford Nanopore Technologies MinION FLO-MIN106 flow cell from a library prepared with the SQK-RAD004 kit using 1 μg of DNA and 20-s tagmentation. Base calling was performed using Guppy v3.1.5. Sequence quality was assessed using FastQC v0.11.9 (5) for Illumina reads and MinIONQC.R v1.4.2 (6) for Nanopore reads. The preprocessing of Illumina sequencing data using Trimmomatic v0.39 (7) (parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50) resulted in 3,989,730 paired-end reads (mean lengths of 223 bp and 164 bp for forward and reverse reads, respectively, with ~200× coverage). Nanopore long reads (total of 2.2 Gbp, with an  $N_{50}$  value of 33,388 bp) were filtered and trimmed for length (>1,500 bp) and quality (scores of >8) using NanoFilt v2.5.0 (8), and adapters were removed using Porechop v0.2.4 (9). Long reads were further reduced to 800 Mbp as a target quantity using FilTlong

**Citation** Oger-Desfeux C, Briolay J, Oger PM, Lafay B. 2021. Complete genome sequence of *Bradyrhizobium* sp. strain BDV5419, representative of Australian genospecies L. Microbiol Resour Announc 10:e00042-21. <https://doi.org/10.1128/MRA.00042-21>.

**Editor** John J. Dennehy, Queens College

**Copyright** © 2021 Oger-Desfeux et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Bénédicte Lafay, [benedicte.lafay@univ-lyon1.fr](mailto:benedicte.lafay@univ-lyon1.fr).

**Received** 13 January 2021

**Accepted** 3 February 2021

**Published** 25 February 2021

v0.2.0 (10) (parameters: --min\_length 2000 --keep\_percent 90 --target\_bases 800000000). Illumina and Nanopore reads were coassembled using Unicycler v.0.4.8 (11) with default parameters. This resulted in a single circular chromosome of length 7,401,610 bp, rotated to start at *dnaA*. Its average G+C content is 64.6%. The assembly was carefully inspected by visualizing the alignment of long and short reads using minimap2 v2.17 (12) and IGV v2.7.2 (13). The genome sequence was automatically annotated by the NCBI Prokaryote Genome Annotation Pipeline (PGAP) v4.13 (14). The genome consists of 6,879 protein-coding genes, 48 tRNAs, 1 copy each of the 5S, 16S, and 23S rRNA genes, and 94 pseudogenes.

**Data availability.** The genome sequence of *Bradyrhizobium* genospecies L strain BDV5419 is available in NCBI GenBank under accession number [CP061378](https://doi.org/10.1093/nar/gkw569). The raw sequence reads are available under SRA accession numbers [SRX9514897](https://www.ncbi.nlm.nih.gov/sra/SRX9514897) and [SRX9514899](https://www.ncbi.nlm.nih.gov/sra/SRX9514899) under BioProject number [PRJNA662585](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA662585) and BioSample number [SAMN16089660](https://www.ncbi.nlm.nih.gov/biosample/SAMN16089660).

## ACKNOWLEDGMENT

This project was supported by the École Centrale de Lyon BQR Program.

## REFERENCES

- Lafay B, Burdon JJ. 1998. Molecular diversity of rhizobia occurring on native shrubby legumes in southeastern Australia. *Appl Environ Microbiol* 64:3989–3997. <https://doi.org/10.1128/AEM.64.10.3989-3997.1998>.
- Vincent JM. 1970. A manual for the practical study of root-nodule bacteria. International Biological Programme handbook 15. Blackwell Science Publications, Oxford, England.
- Dhaese P, De Greve H, Decraemer H, Schell J, Van Montagu M. 1979. Rapid mapping of transposon insertion and deletion mutations in the large T-plasmids of *Agrobacterium tumefaciens*. *Nucleic Acids Res* 7:1837–1849. <https://doi.org/10.1093/nar/7.7.1837>.
- Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. <https://github.com/najoshi/sickle>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Lanfear R, Schalamun M, Kainer D, Wang W, Schwessinger B. 2019. MinIONQC: fast and simple quality control for MinION sequencing data. *Bioinformatics* 35:523–525. <https://doi.org/10.1093/bioinformatics/bty654>.
- 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>.
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
- Wick RR. 2018. Porechop. <https://github.com/rwick/Porechop>.
- Wick RR. 2018. Filtlong. <https://github.com/rwick/Filtlong>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Li H. 2018. minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative Genomics Viewer. *Nat Biotechnol* 29:24–26. <https://doi.org/10.1038/nbt.1754>.
- Tatusova T, DiCuccio M, Badredin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.