





Contiguous Genome Sequence of *Frankia* sp. Strain Arl3, Isolated from Root Nodules of *Alnus rubra* Bong

 Callum J. Bell,^a Johnny A. Sena,^a Isaac S. Gifford,^b  Alison M. Berry^b

^aNational Center for Genome Resources, Santa Fe, New Mexico, USA

^bDepartment of Plant Sciences, University of California, Davis, California, USA

ABSTRACT We report the genome sequence of *Frankia* sp. strain Arl3, recovered as a single contig from one run of the Oxford Nanopore Technologies (ONT) MinION instrument. The genome has a G+C content of 72%, is 7,541,222 bp long, and contains 5,427 predicted protein-coding genes.

We sequenced the genome of *Frankia* sp. strain Arl3, a nitrogen-fixing taxon (phylum *Actinobacteria*) isolated from the root nodules of *Alnus rubra* plants grown in a greenhouse at Harvard Forest (Petersham, MA, USA) (1). The root nodules were initiated using a crushed-nodule inoculum derived from field-grown nodules of *Alnus rubra*, provided by L. E. Baribo (Weyerhaeuser Co., Tacoma, WA). Despite having been isolated in 1978 (1) and widely used experimentally (2–8), Arl3 has not been sequenced to date.

Arl3 was grown in basal medium with propionate (BAP) (9) supplemented with 5 mM pyruvate and 5 mM morpholinepropanesulfonic acid (MOPS), as in reference 10. Hyphae were collected using a 14G syringe needle, centrifuged at 10,000 rpm for 5 min, resuspended, and homogenized by passage through a 21G needle. The homogenate was inoculated into fresh medium and grown for 1 week at 28°C with shaking at 50 rpm. Hyphae were collected as above, then resuspended in Tris-EDTA (TE) buffer.

Genomic DNA (gDNA) was extracted using the cetyltrimethylammonium bromide (CTAB) protocol in reference 10, purified by phenol-chloroform extraction, and stored in TE buffer at –20°C. The DNA was quantified using the Qubit 3.0 fluorometer with the double-stranded DNA (dsDNA) broad-range (BR) assay kit, cleaned and concentrated by mixing with 0.45× volume AMPure beads (Beckman Coulter Life Sciences) according to the manufacturer's instructions, and used for library preparation according to the ONT 1D gDNA protocol (SQK-LSK109).

Sequencing was performed on a MinION device (MinKNOW version 3.3.2) connected to a MinIT processing unit. Guppy version 3.0.3 (11) was used for real-time base calling in high-accuracy mode, with a quality score cutoff of 7. We obtained 673,563 reads averaging 4,608 bp, with a maximum length of 109,884 bp.

The sequence reads in FASTQ format were assembled using Flye version 2.7.1 (12), specifying the expected genome size, and aligned back to the draft assembly using Minimap2 version 2.17 (13). After conversion to BAM format (14), the alignments were assessed for quality and read alignment rate using PycoQC version 2.5.0.21 (15). The assembly was polished using Racon version v1.4.16 (16), followed by a second polishing step using Medaka version 1.0.1 (<https://nanoporetech.github.io/medaka/>).

The genome was recovered as a single contig 7,541,222 bp long, with a 72% G+C content. In total, 3,104,208,283 nucleotides were sequenced, representing 411× coverage. Annotation of the corresponding RefSeq record at NCBI with the Prokaryotic Genome Annotation Pipeline (PGAP) (17) predicted 5,427 protein-coding genes. CheckM version 1.1.3 (18) indicated that 98% of the *Actinomycetales* marker genes

Citation Bell CJ, Sena JA, Gifford IS, Berry AM. 2021. Contiguous genome sequence of *Frankia* sp. strain Arl3, isolated from root nodules of *Alnus rubra* Bong. *Microbiol Resour Announc* 10:e00800-21. <https://doi.org/10.1128/MRA.00800-21>.

Editor Leighton Pritchard, SIPBS, University of Strathclyde

Copyright © 2021 Bell 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 Callum J. Bell, cjb@ncgr.org.

Received 12 August 2021

Accepted 27 September 2021

Published 28 October 2021

were present, with zero contamination or strain heterogeneity. Arl3 is closely related to Cpl1-P (19) (GenBank accession number [GCA_001421075.1](https://www.ncbi.nlm.nih.gov/nuclseq/GCA_001421075.1)) and Cpl1-S (19) ([GCA_000948395.1](https://www.ncbi.nlm.nih.gov/nuclseq/GCA_000948395.1)), sharing digital DNA-DNA hybridization (dDDH) (20) values of 99.3% and 99.5%, respectively. Plasmids previously reported in strain Arl3 (21) were not detected in this study. Underrepresentation of plasmids in ONT sequencing has been noted previously (22). Unless noted otherwise, all software was run using default parameters.

Data availability. The raw data and assembled genome sequence are indexed under NCBI BioProject accession number [PRJNA747286](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA747286). The assembled genome sequence has been assigned GenBank accession number [CP079862](https://www.ncbi.nlm.nih.gov/nuclseq/CP079862). The MinION reads have been assigned the Sequence Read Archive accession number [SRR15181822](https://www.ncbi.nlm.nih.gov/sra/SRR15181822).

ACKNOWLEDGMENTS

This work was supported in part by NSF Plant Genome Research Program award 1547842 (C.J.B.), UC Davis Plant Sciences Graduate fellowship (I.S.G.), and USDA-NIFA-CA-D-PLS-2173-H (A.M.B.).

REFERENCES

- Berry A, Torrey JG. 1979. Isolation and characterization in vivo and in vitro of an actinomycetous endophyte from *Alnus rubra* Bong, p 69–83. *In* Gordon JC, Wheeler CT, Perry DA (ed), Symbiotic nitrogen fixation in the management of temperate forests. Forest Research Laboratory, Oregon State University, Corvallis, OR.
- Fontaine MS, Young PH, Torrey JG. 1986. Effects of long-term preservation of Frankia strains on infectivity, effectivity, and in vitro nitrogenase activity. *Appl Environ Microbiol* 51:694–698. <https://doi.org/10.1128/aem.51.4.694-698.1986>.
- Stevens GA, Jr, Berry AM. 1988. Cytokinin secretion by Frankia sp. HFP Arl3 in defined medium. *Plant Physiol* 87:15–16. <https://doi.org/10.1104/pp.87.1.15>.
- Murry MA, Lopez MF. 1989. Interaction between hydrogenase, nitrogenase, and respiratory activities in a Frankia isolate from *Alnus rubra*. *Can J Microbiol* 35:636–641. <https://doi.org/10.1139/m89-102>.
- Mirza BS, Welsh A, Hahn D. 2007. Saprophytic growth of inoculated Frankia sp. in soil microcosms. *FEMS Microbiol Ecol* 62:280–289. <https://doi.org/10.1111/j.1574-6941.2007.00382.x>.
- Leul M, Normand P, Sellstedt A. 2009. The phylogeny of uptake hydrogenases in Frankia. *Int Microbiol* 12:23–28.
- Gabbarini LA, Wall LG. 2011. Diffusible factors involved in early interactions of actinorhizal symbiosis are modulated by the host plant but are not enough to break the host range barrier. *Funct Plant Biol* 38:671–681. <https://doi.org/10.1071/FP11003>.
- Samant S, Huo T, Dawson JO, Hahn D. 2016. Abundance and relative distribution of Frankia host infection groups under actinorhizal *Alnus glutinosa* and non-actinorhizal *Betula nigra* trees. *Microb Ecol* 71:473–481. <https://doi.org/10.1007/s00248-015-0643-2>.
- Murry MA, Fontaine MS, Torrey JG. 1984. Growth kinetics and nitrogenase induction in Frankia sp. HFP Arl 3 grown in batch culture. *Plant Soil* 78: 61–78. <https://doi.org/10.1007/BF02277840>.
- Gifford I, Vance S, Nguyen G, Berry AM. 2019. A stable genetic transformation system and implications of the type IV restriction system in the nitrogen-fixing plant endosymbiont Frankia alni ACN14a. *Front Microbiol* 10: 2230. <https://doi.org/10.3389/fmicb.2019.02230>.
- Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Leger A, Leonardi T. 2019. pycoQC, interactive quality control for Oxford Nanopore sequencing. *J Open Source Softw* 4:1236. <https://doi.org/10.21105/joss.01236>.
- Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746. <https://doi.org/10.1101/gr.214270.116>.
- Tatusova T, DiCuccio M, Badretdin 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>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Oshone R, Hurst SG, IV, Abebe-Akele F, Simpson S, Morris K, Thomas WK, Tisa LS. 2016. Permanent draft genome sequences for two variants of Frankia sp. strain Cpl1, the first Frankia strain isolated from root nodules of *Comptonia peregrina*. *Genome Announc* 4:e01588-15. <https://doi.org/10.1128/genomeA.01588-15>.
- Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10: 2182. <https://doi.org/10.1038/s41467-019-10210-3>.
- Normand P, Simonet P, Butour JL, Rosenberg C, Moiroud A, Lalonde M. 1983. Plasmids in Frankia sp. *J Bacteriol* 155:32–35. <https://doi.org/10.1128/jb.155.1.32-35.1983>.
- Wick RR, Judd LM, Wyres KL, Holt KE. 2021. Recovery of small plasmid sequences via Oxford Nanopore sequencing. *Microb Genom* 7. <https://doi.org/10.1099/mgen.0.000631>.