AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Agrobacterium fabrum* Strain 1D1104

Naxin Huo,[a] Yong Gu,[a] Kent F. McCue,[a] Diaa Alabed,[a] James G. Thomson[a]

[a]USDA-ARS Crop Improvement and Genetics, Western Regional Research Center, Albany, California, USA

ABSTRACT This work reports the draft genome sequence of *Agrobacterium fabrum* strain 1D1104. The assembled genome is composed of a 2,774,783-bp circular chromosome, a 2,110,112-bp linear chromosome, an AT plasmid of 133,577 bp, and four unassembled contigs of 5,389,544 bp, 42,391 bp, 41,768 bp, and 35,476 bp.

Here, we present the novel genome sequence of *Agrobacterium fabrum* strain 1D1104, obtained from C. I. Kado's retired microbe collection at UC Davis. The collection is being maintained by G. Coaker at UC Davis. Strain 1D1104 was obtained from a poplar tree gall by the Kobe lab. On 15 March 1972, the isolated culture was lyophilized in a glass tube for long-term storage. Further characterization by the Kobe lab indicates that the isolated strain was designated an agrobacterium, based on a positive 3-keto-lactose test result, but did not determine if it is pathogenic (gall forming). The lyophilized strain was revived and grown in Luria broth at 28 to 30°C, with shaking at 200 rpm.

Genomic DNA was isolated from 1D1104 (1) using the Qiagen blood and cell culture DNA maxi kit (number 13362) and genomic DNA buffer set (number 19060) (Qiagen) (2). The DNA sample was evaluated by gel electrophoresis and quantified using both a 2100 NanoDrop spectrophotometer (Thermo Fisher Scientific) and a Qubit fluorimeter (Invitrogen) with the Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit (Invitrogen). The genomic DNA was sheared using a g-TUBE device (Covaris). A 20-kb DNA library was constructed according to the manufacturer's instructions, size selected using the BluePippin system, and sequenced using single-molecule real-time (SMRT) sequencing technology on a PacBio RS II system. SMRT sequencing data were generated at an average coverage of 87.21×, with a mean read length of 18,403 bp. *De novo* genome assembly was conducted with 26,553 sequence reads using the Hierarchical Genome Assembly Process (HGAP) workflow with the default parameter value and the genome size set at 5 Mb (SMRT Portal; Pacific Biosciences), the RS_HGAP_Assembly.3 protocol (3), and Smrtanalysis_2.3.0 software (https://www.pacb.com/wp-content/uploads/SMRT-Analysis-Software-Installation-v2-3-0.pdf).

This allowed the generation of 7 polished contigs, with a contig N_{50} length of 2,798,919 bp; the sum of the contig lengths was 5,216,137 bp. The sequence was annotated, and a chimeric overlap of 24,128 bp was identified and removed. The duplicate region was determined to be a sequence overlap of the circular chromosome, which has a final size of 2,774,783 bp and a GC content of 59.7%. The linear chromosome was determined to be 2,110,112 bp with a GC content of 59.7%. The predicted AT plasmid was composed of 133,577 bp with a GC content of 58.6%. Four contigs that could not be aligned to the genome were also observed and determined to comprise 53,895 bp (P53895) with a GC content of 55.3%, 42,391 bp (P42391) with a GC content of 56.5%, 41,768 bp (P41768) with a GC content of 57.2%, and 35,476 bp (P35476) with a GC content of 56.5%. No virulence plasmid (pTi) was detected during sequence analysis.

The assembled and raw read sequences were entered into the National Center for Biotechnology Information (NCBI), and BLAST was used for identification (http://blast.ncbi.nlm.nih.gov/). Automated annotation was performed with the Rapid Annotations using

Editor Julia A. Maresca, University of Delaware

Address correspondence to James G. Thomson, James.Thomson@usda.gov.

Received 6 October 2021
Accepted 5 November 2021
Published 2 December 2021

Subsystem Technology (RAST) pipeline for annotation of the genome (4). *Agrobacterium fabrum* strain 1D1104 contains 5,083 predicted coding sequences, 542 subsystems, and 67 predicted RNA-coding genes, as curated by SEED data (http://theseed.org). Based on a RAST analysis, genomic comparison shows the 1D1104 circular chromosome to be related to *Agrobacterium tumefaciens* strain CCNWGS0286, while the assembled 1D1104 linear chromosome appears related to *Agrobacterium tumefaciens* strain 12D1.

Data availability. The whole-genome assembly for *Agrobacterium fabrum* ("*Agrobacterium tumefaciens*") strain 1D1104 has been deposited in DDBJ/ENA/GenBank under the BioProject accession number PRJNA750239, BioSample accession number SAMN20446794, and SRA accession number SRS9645414.

ACKNOWLEDGMENT

This work was supported by USDA Agricultural Research Service CRIS projects 2030-21220-002-00-D and 2030-21430-014-00-D. Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

REFERENCES

1. Wise AA, Liu Z, Binns AN. 2006. Nucleic acid extraction from *Agrobacterium* strains, p 67–76. *In* Wang K (ed), Agrobacterium protocols, 2nd ed, vol 1. Humana Press, Totowa, NJ. https://doi.org/10.1385/1-59745-130-4:67.
2. Qiagen. CLC Genomics Workbench. 8.5. Qiagen, Redwood City, CA. https://www.qiagenbioinformatics.com/.
3. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
4. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.