




Draft Genome Sequence of *Agrobacterium tumefaciens* Strain 1D1526

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

^aUSDA-ARS Crop Improvement and Genetics, Western Regional Research Center, Albany, California, USA

ABSTRACT This work reports the draft genome sequence of *Agrobacterium tumefaciens* strain 1D1526. The assembled genome is composed of a 2,881,823-bp circular chromosome, a 2,235,711-bp linear chromosome, and a 44,582-bp unassembled contig.

Here, we present the novel genome from *Agrobacterium tumefaciens* strain 1D1526, obtained from the Kobe microbe collection at UC Davis. Strain 1D1526 was obtained from an apple tree gall by the Kobe lab on 11 July 1982. There was no further characterization by the Kobe lab to indicate whether the isolated strain was pathogenic (gall forming). The strain was grown in Luria broth at 28 to 30°C with shaking at 200 rpm.

Genomic DNA was isolated from our strains (1) using the Qiagen blood and cell culture DNA maxikit (catalog number 13362) and genomic DNA buffer set (catalog number 19060) (2). DNA samples were evaluated using gel electrophoresis and quantified using both a 2100 Nanodrop spectrophotometer (Thermo Fisher Scientific) and a Qubit fluorometer (Invitrogen) with the Qubit double-stranded DNA (dsDNA) HS assay kit (Invitrogen). The genomic DNA was sheared with a g-TUBE (Covaris). A 20-kb DNA library was constructed according to the manufacturer's instructions using the Blue-Pippin size selection system and sequenced using single-molecule real-time (SMRT) sequencing technology on a PacBio RS system. SMRT sequencing data were generated at an average coverage of $87.21 \times$ 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 using the default parameter value and a genome size set at 5 Mb (SMRT Portal; Pacific Biosciences), protocol RS_HGAP_Assembly.3 (3), and SMRTAnalysis_2.3.0 software (<https://www.pacb.com/wp-content/uploads/2015/09/SMRT-Analysis-Software-Installation-v2.3.0.pdf>).

This allowed the generation of 3 polished contigs with an N_{50} contig length of 2,907,945 bp and a sum of contig lengths of 5,188,238 bp. The DNA was manually circularized via chimeric overlap of 26,082 bp for the circular chromosome, which has a final composition of 2,881,823 bp with a GC content of 59.6%. The linear chromosome was determined to be 2,235,711 bp with a GC content of 59.6%. An unincorporated DNA contig was also observed and determined to be 44,582 bp with a GC content of 58.4%. Neither an *A. tumefaciens* plasmid nor a virulence Ti tumor-inducing (Ti) plasmid vector was detected during sequence analysis.

Assembled and raw sequence reads 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 using the Rapid Annotation using Subsystem Technology (RAST) pipeline for annotation of the genome (4). *Agrobacterium tumefaciens* strain 1D1526 contains 4,935 predicted coding sequences, 536 subsystems, and 65 predicted RNA-coding genes as curated by SEED data (<http://theSEED.org>). Based on RAST analysis, genomic comparison shows the 1D1526 circular chromosome to be related to *Agrobacterium tumefaciens* strain CCNWGS0286, while the

Citation Huo N, Gu Y, McCue KF, Alabed D, Thomson JG. 2019. Draft genome sequence of *Agrobacterium tumefaciens* strain 1D1526. *Microbiol Resour Announc* 8:e01084-19. <https://doi.org/10.1128/MRA.01084-19>.

Editor Christina A. Cuomo, Broad Institute
This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.
Address correspondence to James G. Thomson, James.Thomson@ars.usda.gov.

Received 13 September 2019

Accepted 16 October 2019

Published 7 November 2019

assembled 1D1526 linear chromosome appears to be related to *Agrobacterium tumefaciens* strain C58. With MacVector version 17.0.0 DNA matrix analysis, the assembled contigs were compared to the predicted C58 circular and linear chromosome sequences (GenBank accession numbers [AE007869](#) and [AE007870](#), respectively). The 1D1526 and C58 circular chromosome sequences are syntenic for their entire lengths, with ~94% exhibiting 95% or higher identity and with minor deletions of 28 and 35 kb and two insertions of 45 and 75 kb relative to C58. The 1D1526 linear chromosome is syntenic for its entire length, with ~80% exhibiting 95% or higher identity with more than a dozen insertions/deletions. The unincorporated DNA contig appears to be a portion of the linear chromosome as based on the NCBI nucleotide BLAST database and provides 88% identity (53% coverage) to the *Agrobacterium tumefaciens* strain CFBP6625 linear chromosome in the region coincident with a 30-kb insertion in CFBP6625, a relative of C58.

Data availability. The whole-genome assembly for *Agrobacterium tumefaciens* strain 1D1526 has been deposited in DDBJ/ENA/GenBank under the BioProject accession number [PRJNA546276](#), BioSample accession number [SAMN11958781](#), SRA accession number [SRP201107](#), and whole-genome sequence (WGS) accession number [VTZQ00000000](#).

ACKNOWLEDGMENTS

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. 2019. CLC Genomics Workbench. 8.5. Qiagen, Redwood City, CA. <https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/>.
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. Non-hybrid, 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>.