AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of a Gram-Positive Bacterium, *Leifsonia* sp. Strain PS1209, a Potato Endophyte

Yingyu Liu,[a] Xing Ma,[a] Tyler C. Helmann,[b] Heather McLane,[a] Paul Stodghill,[a,b] Bryan Swingle,[a,b] Melanie J. Filiatrault,[a,b] Keith L. Perry[a]

[a]School of Integrative Plant Science, Section of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, New York, USA

[b]Emerging Pests and Pathogens Research Unit, Robert W. Holley Center for Agriculture and Health, Agricultural Research Service, U.S. Department of Agriculture, Ithaca, New York, USA

ABSTRACT We report the complete and annotated genome sequence of a Gram-positive bacterium, *Leifsonia* sp. strain PS1209, a potato endophyte that was isolated from apparently healthy tubers of potato cultivar NY166. The circular genome is 4,091,164 bp long, with a GC content of 69.08%, containing 3,926 genes.

In 2019, we isolated a Gram-positive endophyte from apparently healthy tubers of potato cultivar NY166 (potato breeding program, Cornell University). This endophyte persisted in bud tissues following surface sterilization (10% commercial bleach). It was isolated at room temperature after being grown from intact bud tissue using nutrient broth (1) and Richardson's solution (2) after 2 to 3 days of incubation. The sequence of the 16S rRNA gene (3) amplified from genomic DNA showed 99.9% identity to the 16S rRNA gene of *Leifsonia lichenia* 2Sb^T (4).

A single-colony isolate of *Leifsonia* sp. PS1209 was stored in 20% glycerol at −80°C and cultured in Luria-Bertani medium (5) at 28°C. High-molecular-weight DNA was extracted from overnight bacterial cultures (after approximately three passages after isolation) using a Wizard genomic DNA purification kit (Promega, USA). DNA quantity and integrity were assessed using a NanoDrop 1000 spectrophotometer and Qubit 3.0 fluorometer (Thermo Fisher Scientific, USA) and a 2100 Bioanalyzer system (Agilent, USA), respectively.

For long-read sequencing, the DNA library was prepared using the rapid sequencing kit SQK-RAD004 on an R9.4.1 SpotON FLO-MIN106 flow cell and was sequenced with a MinION Mk1B device. Base calling was performed with Guppy v3.5.1. This generated 376,226 reads with a total length of 2,465,189,756 bp, an N_{50} value of 13,374 bp, and genome coverage of 598×. The genome was *de novo* assembled using Flye v2.5 (6) and polished with Nanopolish v0.12.2a (7). For short-read sequencing, the same DNA as for Nanopore sequencing was used for library preparation. The DNA library was prepared using the NEBNext Ultra II FS DNA library preparation kit for Illumina with NEBNext multiplex oligonucleotides (New England Biolabs, USA). AMPure XP beads (Beckman Coulter, USA) were used for DNA size selection and purification. The DNA library was sequenced on a MiSeq instrument (Illumina, USA) with the 2 × 250-bp mode, yielding 1,877,134 paired-end reads. Duplicate read pairs were removed using a customized Perl script (https://github.com/Sunhh/NGS_data_processing/blob/master/drop_dup_both_end.pl) (8). Deduplicated reads were processed to remove the adaptors, low-quality sequences (Q < 20), and short reads (<50 bp) using Trimmomatic v0.39 (9). All software systems were run with their default settings unless otherwise noted. BWA v0.7.17 (10) and Pilon v1.22 (11) were used to align the resulting 1,590,269 high-quality cleaned Illumina read pairs to the draft genome and to correct the draft

Citation Liu Y, Ma X, Helmann TC, McLane H, Stodghill P, Swingle B, Filiatrault MJ, Perry KL. 2020. Complete genome sequence of a Gram-positive bacterium, *Leifsonia* sp. strain PS1209, a potato endophyte. Microbiol Resour Announc 9:e00447-20. https://doi.org/10.1128/MRA.00447-20.

Editor David A. Baltrus, University of Arizona

Address correspondence to Melanie J. Filiatrault, melanie.filiatrault@usda.gov.

Received 20 April 2020
Accepted 3 June 2020
Published 25 June 2020

genome for three rounds, with a genome coverage of 128×. Circlator v1.5.5 was used with the fixstart argument to set the position number to start at the *dnaA* gene (12).

The circular *Leifsonia* sp. strain PS1209 genome is 4,091,164 bp long and was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.11 (13). It contains 3,926 genes in total, with 3,819 protein-coding genes, 52 pseudogenes, 2 5S rRNA-coding sequences, 2 16S rRNA-coding sequences, 2 23S rRNA-coding sequences, 46 tRNAs, and 3 noncoding RNAs. The GC content of PS1209 of 69.08% is 9% higher than that of 2Sb^T but similar to that of other *Leifsonia* spp. (4). The genome sequence of the type strain 2Sb is not publicly available but is needed to unambiguously determine the species of PS1209.

Data availability. The *Leifsonia* sp. PS1209 genome sequence has been deposited in GenBank under the accession number CP051154. The Nanopore and Illumina raw reads have been deposited in the Sequence Read Archive (SRA) under the accession numbers SRR11498590 and SRR11498591, respectively.

ACKNOWLEDGMENTS

This project was funded by the National Institute for Food and Agriculture, Specialty Crops Research Initiative (grant 2017-51181-26827). The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

We thank the Genomics Facility of the Biotechnology Resource Center at the Cornell University Institute of Biotechnology for help with sequencing experiments.

We declare no conflicts of interest.

Mention of trade names or commercial products in this publication 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. Fahy PC, Persley GJ. 1983. Plant bacterial diseases: a diagnostic guide. Academic Press, Sydney, Australia.
2. Richardson LT. 1957. Quantitative determination of viability of potato ring rot bacteria following storage, heat, and gas treatments. Can J Bot 35:647–656. https://doi.org/10.1139/b57-054.
3. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 1991. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173:697–703. https://doi.org/10.1128/jb.173.2.697-703.1991.
4. An S-Y, Xiao T, Yokota A. 2009. *Leifsonia lichenia* sp. nov., isolated from lichen in Japan. J Gen Appl Microbiol 55:339–343. https://doi.org/10.2323/jgam.55.339.
5. Bertani G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol 62:293–300. https://doi.org/10.1128/JB.62.3.293-300.1951.
6. 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.
7. Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled de novo using only Nanopore sequencing data. Nat Methods 12:733–735. https://doi.org/10.1038/nmeth.3444.
8. Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z, Zhang F, Tian J, Lucas WJ, Doyle JJ, Li H, Fei Z, Xu Y. 2017. Karyotype stability and unbiased fractionation in the paleo-allotetraploid *Cucurbita* genomes. Mol Plant 10:1293–1306. https://doi.org/10.1016/j.molp.2017.09.003.
9. 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.
10. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
11. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
12. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.
13. 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.