AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Applied and Industrial Microbiology | Announcement

Complete genome of *Erwinia spp. str*. LJJL01 isolated from waste charcoal

Lakshika Dissanayake,[1] **Jeffrey G. Linger,**[2] **Lahiru N. Jayakody**[1,3]

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT We present the complete genome sequence of *Erwinia spp. str*. LJJL01, isolated from waste charcoal in Colorado, USA, using Oxford Nanopore sequencing. This sequence provides important insights into this bacterium's metabolic and catabolic robustness to utilize sugars, acids, and aromatics, highlighting its potential as a bio-industrial strain for various feedstocks.

KEYWORDS *Erwinia spp. str*. LJJL01, Oxford Nanopore sequencing, synthetic microbiology

Erwinia spp. str. LJJL01 was isolated from the top 6 inches of a soil sample that contains hardwood lump charcoal in 2016 from Denver, Colorado, USA (Global positioning system coordinates: 39.778447–105.036688), based on its ability to grow on an agar plate containing minimal media (M9) supplemented with 5 mM of lignin-derived monomers such as syringol or catechol (1). Earlier this strain was identified as *Erwinia aphidicola* LJJL01 based on a partial sequence of 16S rRNA. The full-length 16S rRNA gene sequence based on the assembled genome reported here and phylogenetic analysis suggests that this bacterium is another species of *Erwinia* and is designated as *Erwinia spp. str*. LJJL01 (Fig. 1) (2, 3). The strain was deposited in the U.S. Department of Agriculture's Agricultural Research Service Culture Collection under the number NRRL B-65680. The first draft genome sequence of *Erwinia spp. str*. LJJL01 was completed by the Joint Genome Institute in 2018 (ID- 2751185488) using PacBio RS II, sequencing with 122.0× genome coverage and assembly with HGAP (version 2.3.0_p5) and annotated using IMG Annotation Pipeline (version 4.15.1; NCBI GenBank ID: SNXL01000001:SNXL01000004) but was not described in peer-reviewed literature.

Here, we present the re-sequenced and closed genome of *Erwinia spp. str*. LJJL01 using nanopore sequencing. A single colony of *Erwinia spp. str*. LJJL01 from a Luria-Bertani agar plate was used to inoculate 5 mL of M9-glucose medium, incubated overnight at 30°C and 225 rpm to obtain cell pellets for genomic DNA extraction. DNA was extracted using GeneJET genomic DNA purification kit following the manufacturer's instructions (Thermo Fisher Scientific, Waltham, MA, USA). DNA concentration was measured using nanoquant plate in TECAN Sunrise Infinite plate reader (TECAN, Switzerland). Genome sequencing was performed using Oxford Nanopore Technology (ONT) at Eurofins Genomics LLC (Louisville, KY, USA). DNA was prepared for sequencing using the Rapid Barcoding Kit 96 V14 SQK-RBK114.96 following the manufacturer's instructions (ONTs, UK). Mechanical shearing or size selection of the library was not adopted in the analysis. Sequencing was performed using Oxford Nanopore GridION loaded with V14 MinION flow cell (Oxford Nanopore Technologies, UK). The base calling was carried out using Dorado (version 4.3) (4). The resulting reads undergo quality filtering using NanoPack2 software, are assembled with Flye assembler (version 2.9.2), polished with Pilon (version 1.24), and the assembly quality is assessed using QUAST (version 5.2.0). The default parameters were used for the analysis (5–7). A summary of the

Editor Elinne Becket, California State University San Marcos, San Marcos, California, USA

Address correspondence to Lahiru N. Jayakody, lahiru.jayakody@siu.edu.

The authors declare no conflict of interest.

Received 5 April 2025
Accepted 26 July 2025
Published 15 August 2025

[This article was published on 15 August 2025 with an error in the author byline. The byline was corrected in the current version, posted on 21 August 2025.]

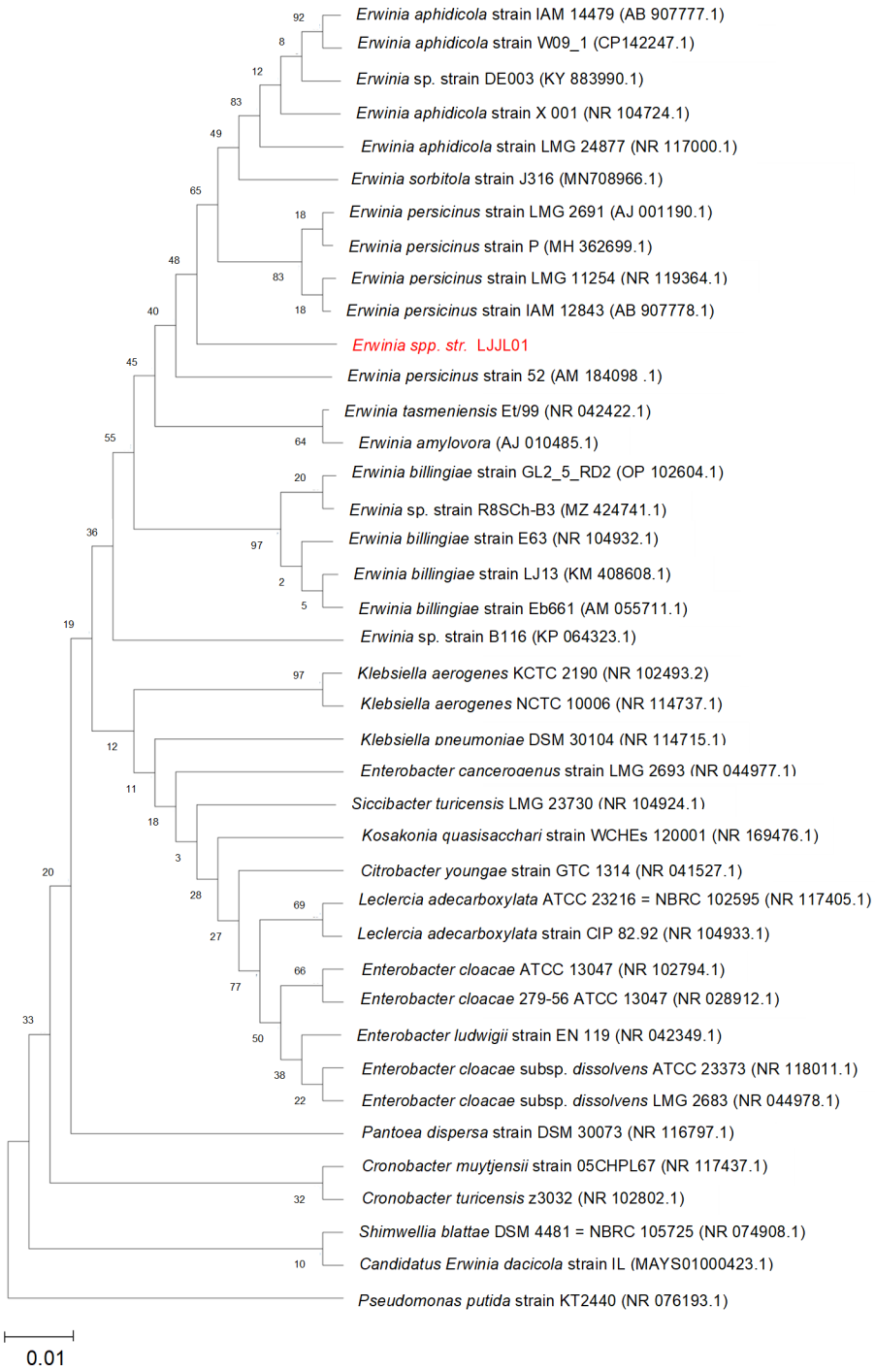

FIG 1 Maximum likelihood phylogenetic tree based on 16S rRNA sequences of 35 species aligned with the test strain. Full-length 16S rRNA genes aligned by SILVA (version 1.2.11) and unrooted maximum likelihood phylogenetic tree was generated by MEGA X, with 1,000 bootstraps. The scale bar represents 1% identity dissimilarity between sequences.

results is provided in Table 1. Additionally, the assembly is annotated using the NCBI Prokaryotic Annotation Pipeline (version 6.7) and available under GCA_047300805.1 (8–10).

Nanopore genome sequencing of *Erwinia spp. str.* LJJL01 revealed the presence of four contigs, including the genome of size 4,751,215 bp, and three native plasmids of sizes 104,322 bp, 89,813 bp, and 5973 bp. Of note, chromosomes and plasmids

TABLE 1 Nanopore sequencing statistics and genomic features of *Erwinia spp. str.* LJJL01[a]

Overall assembly	
Total assembly length (bp)	4,951,323
No. of contigs	4
Largest contig	4,751,215
GC %	56.56
Total reads	145,853
Total bases (Mb)	464,063,254
Average coverage depth	94
N50	4,751,215
L50	1
Assembly completeness	
Marker lineage	Enterobacteriaceae (UID5054)
Completeness	99.13%
Contamination	0.20%
Species determination	
Best species match	*Candidatus Erwinia dacicola* strain IL
Sequence identity to best species match	92.5557%
Summary of annotations	
tRNAs	83
rRNAs	22
ncRNAs	10
CDSs	4,510
Pseudogenes	110

[a]CDS, Coding DNA Sequence; GC, guanine and cytosine; ncRNAs, Non-coding RNA.

were confirmed as circular using Flye (6). The genome includes genes that encode enzymes for metabolic pathways involving sugars (glucose, xylose, arabinose, galactose, and maltose), acids (lactic and acetic), and aromatics (such as catechol, vanillin, and benzoate).

ACKNOWLEDGMENTS

The National Science Foundation, NSF–2339617-CAREER: Development of novel biocatalyst for consolidated bioprocessing of plastic (CBPP) grant supported the research reported in this publication.

AUTHOR AFFILIATIONS

[1]School of Biological Science, Southern Illinois University Carbondale, Carbondale, Illinois, USA
[2]National Renewable Energy Laboratory, Golden, Colorado, USA
[3]Fermentation Science Institute, Southern Illinois University Carbondale, Carbondale, Illinois, USA

AUTHOR ORCIDs

Lakshika Dissanayake http://orcid.org/0000-0001-8985-021X
Jeffrey G. Linger http://orcid.org/0000-0003-3935-751X
Lahiru N. Jayakody http://orcid.org/0000-0001-8405-5022

AUTHOR CONTRIBUTIONS

Lakshika Dissanayake, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review and editing | Jeffrey G. Linger, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Lahiru N. Jayakody, Conceptualization, Data curation, Formal analysis, Funding

acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The complete genome for *Erwinia spp. str.* LJJL01 has been deposited in NCBI under BioProject PRJNA1208068, including the ONT base-called FASTQ file, SRR32982199. The annotated genome (FASTA) is listed under the NCBI accession number CP178576.1. Native plasmids are available under accession numbers CP178574.1, CP178575.1, and CP192608.1.

REFERENCES

1. Dissanayake L. 2021. Engineering *Erwinia aphidicola* LJJL01 - a catabolic powerhouse to deconstruct and upcycle polyethylene terephthalate, p 206. Southern Illinois University at Carbondale, United States—Illinois.
2. Glöckner FO, Yilmaz P, Quast C, Gerken J, Beccati A, Ciuprina A, Bruns G, Yarza P, Peplies J, Westram R, Ludwig W. 2017. 25 years of serving the community with ribosomal RNA gene reference databases and tools. J Biotechnol 261:169–176. https://doi.org/10.1016/j.jbiotec.2017.06.1198
3. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096
4. Prior K, Becker K, Brandt C, Cabal Rosel A, Dabernig-Heinz J, Kohler C, Lohde M, Ruppitsch W, Schuler F, Wagner GE, Mellmann A. 2025. Accurate and reproducible whole-genome genotyping for bacterial genomic surveillance with nanopore sequencing data. J Clin Microbiol 63:2025. https://doi.org/10.1128/jcm.00369-25
5. De Coster W, Rademakers R. 2023. NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics 39:btad311. https://doi.org/10.1093/bioinformatics/btad311
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. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086
8. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105
9. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068
10. 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