

Complete genome sequences of *Pseudomonas allii*, pathogens causing bacterial rot in onion, tomato, Chinese yam, and lettuce in Japan

Yusuke Takashima,¹ Ken Naito,¹ Mamoru Satou,¹ Hiroyuki Sawada¹

AUTHOR AFFILIATION See affiliation list on p. 3.

ABSTRACT Here, we report the complete genome sequences of four strains of *Pseudomonas allii* (MAFF 301514, MAFF 302050, MAFF 730042, and MAFF 730146) that represent pathogens causing bacterial rot in onion, lettuce, Chinese yam, and tomato, respectively, in Japan.

KEYWORDS *Pseudomonas*, plant pathogens

Bacteria-induced rot symptoms have been a challenge to onion cultivation in Japan, and *Pseudomonas marginalis* is considered as a causal agent (1). Previously, we re-identified *P. marginalis* strains obtained from onions as a novel species, *Pseudomonas allii* (2). During quality control in the Genbank Project of National Agriculture and Food Research Organization, three strains obtained from various crops (3–5, Table 1) showed nearly identical 16S rRNA and *rpoD* sequences to the type strain MAFF 301514. It is important to consider whether *P. allii* is a multi-host pathogen or an aggregate of heterogeneous strains, each with a limited host range (5). Here, we present the genome sequences of strains related to *P. allii*, including that of the type strain.

The strains were isolated from Honshu Island in Japan, including prefectures located at the northern (Aomori) and western (Yamaguchi) ends (Table 1), and were preserved under the Genbank Project as a lyophilized state. The strains were cultured on standard methods agar (Nissui) for 3 or 4 days at 28°C. Genomic DNA was extracted using the NucleoBond HMW DNA kit (Macherey-Nagel). Intact DNA was used for the library preparations.

A long-read sequencing library was prepared using the Rapid Barcoding Kit (SQK-RBK004, Oxford Nanopore Technologies [ONT]) and sequenced using PromethION with an R9.4.1 flow cell (FLO-PRO002, ONT). Raw ONT reads (Table 1) were filtered with >10 Phred quality score cutoff by Chopper version 0.5.0 (6). The filtered reads were assembled using Flye version 2.9.2 (7) with default options except for “-g 7 m --asm-coverage 50,” and the assembled circular genomes were polished using Medaka version 1.8.0 (<https://github.com/nanoporetech/medaka>) with the “r941_prom_hac_g507” model using the filtered reads.

Short-read sequencing libraries were prepared using Illumina DNA Prep (M) Tagmentation (Illumina), including the purification and size selection (ca. 500 bp), and 151 bp paired-end reads were sequenced using NovaSeq 6000 (Illumina) (Table 1). Adapter trimming and quality filtering were performed using fastp version 0.23.4 (8) with default settings except for the minimum length set to 151 and adapters specified by a sequencing company. Error corrections consisted of mapping with filtered Illumina reads using bwa-mem2 version 2.2.1 (9) with default except for “-a” option and polishing using Polypolish version 0.5.0 with default settings (10).

Editor David A. Baltrus, The University of Arizona, Tucson, Arizona, USA

Address correspondence to Yusuke Takashima, takashima.yusuke985@naro.go.jp.

The authors declare no conflict of interest.

Received 3 February 2025

Accepted 25 March 2025

Published 15 April 2025

Copyright © 2025 Takashima 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/).

TABLE 1 Sequencing summary and statistics of the genome sequences of four strains of *Pseudomonas allii*

Feature	Culture collection accession no. ^a			
	MAFF 301514	MAFF 302050	MAFF 730042	MAFF 730146
Isolation sources	Onion (<i>Allium cepa</i>)	Lettuce (<i>Lactuca sativa</i> var. <i>capitata</i>)	Chinese yam (<i>Dioscorea polystachya</i> "Nagaimo")	Tomato (<i>Solanum lycopersicum</i>)
Collection location	Hyogo	Tochigi	Aomori	Yamaguchi
Collection year	1980	1978	1977	1993
Original strain names	N-8042	NL 7843	YM 7702	To-1
Isolator	Akira Ohuchi	Yukio Tsuchiya	Manabu Umekawa	Tatsuhiko Karatsu
Sequencing summary				
Nanopore	DRR636248	DRR636249	DRR636250	DRR636251
No. of raw reads ^b /bases (Mbp)	23,920/306	35,212/379	39,707/576	22,714/341
<i>N</i> ₅₀ (bp)	24,261	18,716	25,540	26,573
Illumina	DRR636252	DRR636253	DRR636254	DRR636255
No. of raw reads/bases (Gbp)	3,744,092/0.56	12,388,708/1.85	14,074,772/2.10	10,552,470/1.57
Mean coverage of depth in polypolish	28×	94×	103×	80×
Genome statistics				
Total length (bp)	6,883,462	6,661,385	6,896,138	6,584,391
No. of circular replicons	1	1	1	1
Coverage of depth in Flye	32×	42×	62×	38×
G + C content (mol %)	60.9	60.9	60.8	61.0
No. of CDSs	6,238	6,015	6,274	5,939
Coding ratio (%)	89.7	89.8	89.6	89.8
No. of rRNA operons	5	5	5	5
No. of tRNA	71	73	73	72
CheckM ^c				
(Pseudomonas marker set)				
Completeness (%)	100	100	100	100
Contamination (%)	0.72	0.72	0.72	0.72
OrthoANLu ^d value (%)	–	99.1	98.9	99.1
dDDH ^e value (%)	–	92.6	91.5	92.8
GenBank accession no.	AP039447	AP039448	AP039449	AP039450

^aCulture collection accession numbers (MAFF) under the Genebank Project of National Agriculture and Food Research Organization (https://www.gene.affrc.go.jp/databases-micro_search_en.php).

^bSignals were subjected to basecalling with a high-accuracy model (Guppy version 6.5.7, Oxford Nanopore Technologies [ONT]), demultiplexing, adapter trimming, and low-quality filtering with thresholds such as >1,000 bases and >7 Phred quality scores using MinkNOW version 23.04.5 (ONT).

^cCheckM (14) quality assessment implemented in DFAST was used.

^dOrthoANI using USEARCH. The values were calculated against the type strain (MAFF 301514) of *P. allii* using ANI Calculator on EzBioCloud. "–" indicates that the strain is used as a reference for comparison and no value can be shown.

^edDDH, digital DNA-DNA hybridization. The values were calculated against the type strain (MAFF 301514) of *P. allii* using the Genome-To-Genome Distance Calculator version 3.0 (formula 2) with default options. "–" indicates that the strain is used as a reference for comparison and no value can be shown.

Genomes were annotated using the DDBJ Fast Annotation and Submission Tool (<https://dfast.ddbj.nig.ac.jp/>) with default settings except for the tRNAscan-SE and dnaA rotation function. Table 1 presents the statistics for the genomes determined in this study. OrthoANLu and digital DNA-DNA hybridization (dDDH) values (Table 1), calculated using the ANI Calculator (<https://www.ezbiocloud.net/tools/ani>) (11) and Genome-To-Genome Distance Calculator version 3.0 (<https://ggdc.dsmz.de/ggdc.php#>) (12), respectively, were higher than the Average Nucleotide Identity (ANI) (95%–96%) and dDDH (70%) thresholds for species delineation (13). These results show that the strains are the same species, and the genomes obtained here can be used as a starting point for analyzing intraspecific variations.

ACKNOWLEDGMENTS

This research was carried out under the project “Buildup Breeding Generic Technology for Food Security,” funded by the Ministry of Agriculture, Forestry and Fisheries, Japan. We express our gratitude to Ms. Y. Igaki, Ms. C. Kanazawa, and Ms. H. Nakajima (Research Center of Genetic Resources, National Agriculture and Food Research Organization) for technical assistance. Short-read sequencing using NovaSeq 6000 was conducted by Genome-Lead Co. Ltd. (Takamatsu, Japan).

AUTHOR AFFILIATION

¹Research Center of Genetic Resources, National Agriculture and Food Research Organization (NARO), Tsukuba, Ibaraki, Japan

AUTHOR ORCIDs

Yusuke Takashima  <http://orcid.org/0000-0002-5165-0690>

Ken Naito  <http://orcid.org/0000-0002-0273-7837>

Hiroyuki Sawada  <https://orcid.org/0000-0002-6804-8081>

AUTHOR CONTRIBUTIONS

Yusuke Takashima, Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review and editing | Ken Naito, Methodology, Resources, Writing – review and editing | Mamoru Satou, Resources, Supervision, Writing – review and editing | Hiroyuki Sawada, Conceptualization, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Genome sequences and sequencing reads determined in this study were deposited in DDBJ/ENA/GenBank under Bioproject ID: [PRJDB20186](https://www.ncbi.nlm.nih.gov/bioproject/PRJDB20186). Detailed accession numbers of genome sequences and sequencing reads were listed in Table 1.

REFERENCES

- Ohuchi A, Ohsawa T, Nishimura J. 1983. Two pathogenic bacteria, *Erwinia rhapontici* (Millard 1924) Burkholder 1948 and *Pseudomonas marginalis* pv. *marginalis* (Brown 1918) Stevens 1925, causing a soft rot of onion. *Jpn J Phytopathol* 49:619–626. <https://doi.org/10.3186/jjphytopath.49.619>
- Sawada H, Fujikawa T, Tsuji M, Satou M. 2021. *Pseudomonas allii* sp. nov., a pathogen causing soft rot of onion in Japan. *Int J Syst Evol Microbiol* 71:004582. <https://doi.org/10.1099/ijsem.0.004582>
- Sawada H. 2021. Problems related to *Pseudomonas marginalis* sensu lato. *Plant Protection* 75:135–143.
- Honda M, Kondo T, Tanaka K, Tonouchi A. 2022. Bacterial rot of Chinese yam caused by *Pseudomonas allii*. *J Gen Plant Pathol* 88:325–330. <https://doi.org/10.1007/s10327-022-01073-6>
- Sawada H, Fujikawa T, Satou M. 2023. Dismantling and reorganizing *Pseudomonas marginalis* sensu lato. *Plant Pathol* 72:654–666. <https://doi.org/10.1111/ppa.13690>
- 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>
- 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>
- Chen S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* 2:e107. <https://doi.org/10.1002/imt2.107>
- Vasimuddin M, Misra S, Li H, Aluru S. 2019. Efficient architecture-aware acceleration of BWA-MEM for multicore systems, p 314–324. In 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). Rio de Janeiro, Brazil.
- Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 18:e1009802. <https://doi.org/10.1371/journal.pcbi.1009802>
- Yoon SH, Ha SM, Lim JM, Kwon SJ, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>
- Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. <https://doi.org/10.1186/1471-2105-14-60>
- Riesco R, Trujillo ME. 2024. Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 74:006300. <https://doi.org/10.1099/ijsem.0.006300>
- 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>