

Complete genome sequence of *Pediococcus pentosaceus* TOKAI 10m isolated from Japanese Takuan pickle

Tomoyuki Hibi,^{1,2} Yuki Nakashima,^{3,4} Shin Yasuda,^{1,2,3} Masashi Hirano,^{1,2,3} Hideki Kinoshita,^{1,2,3,5} Takahiro Kawanabe^{2,3}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Pediococcus pentosaceus* TOKAI 10 m, isolated from Japanese Takuan pickle, shows potential as a soy milk yogurt starter. Its genome comprises a 1,788,781 bp chromosome and four plasmids (56,333, 13,649, 10,458, and 10,048 bp), with no virulence factors or antibiotic resistance genes detected.

KEYWORDS complete genome, lactic acid bacteria, *Pediococcus pentosaceus*, Illumina, Oxford Nanopore Technologies

Pediococcus pentosaceus TOKAI 10 m (TK10m), initially identified as *Lactobacillus sakei* MYU 10 (1) and later reclassified and renamed TOKAI 10 m (2). Briefly, 1 g of finely chopped Japanese homemade Takuan pickle was suspended in 10 mL of distilled water, and 0.1 mL of aliquot was plated on De Man, Rogosa, and Sharpe (MRS) (Difco Laboratories, USA) agar and cultured anaerobically at 37°C for 2 days using AnaeroPack™ Anaero (Mitsubishi Gas Chemical, Japan). A single colony was inoculated in MRS broth and cultured aerobically at 37°C for 24 h (1). The 16S rDNA analysis was performed using the 27 f (5'-AGAGTTTGATCCTGGCTCAG-3') and 1525 r (5'-AGAAAGGAGGTGATCCA GCC-3') (3). The homology search using BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) showed high identities (99.32%) with *P. pentosaceus* padder1 (PV916721.1) (2). TK10m can coagulate soy milk (4) and exhibits heavy metal biosorption (1) and functional small RNA production (2). However, since the genes related to the safety and functionality of this strain remained unknown, we performed whole-genome sequencing.

For short-read sequencing, TK10m was cultivated aerobically in 40 mL of MRS broth (Difco Laboratories) at 37°C overnight and harvested via centrifugation. Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (QIAGEN, Germany), and its quality and concentration were assessed using Qubit (Thermo Fisher Scientific, USA). Libraries were prepared using the MiSeq Reagent Kit v3 (Illumina, USA) and sequenced on the Illumina MiSeq platform (PE300). Sequencing produced 2,902,522 reads, totaling 511,484,556 bp. Raw reads were trimmed using fastp v0.24.0 (5) (6) with the parameters “-q 30 -n 20 -t 1 -T 1”.

For long-read sequencing, TK10m was cultured in 5 mL of MRS broth under identical conditions. Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, USA). Quality and concentration were evaluated using NanoDrop 2000c (Thermo Fisher Scientific) and Quantus (Promega). Libraries were prepared using the Rapid Sequencing Kit V14 (SQK-RAD114, Oxford Nanopore Technologies [ONT], UK) without size selection and sequenced on a MinION Mk1B (ONT) employing an R10.4.1 Flow Cell (FLO-FLG114, ONT). Base calling was conducted using Dorado v7.6.8 (high-accuracy model v4.3.0, 400 bps) (<https://github.com/nanoporetech/dorado>). Long-read quality was assessed using NanoPlot v1.44.1 (7), generating 63,812 reads with an N50 of 18,997 bp. Reads were trimmed using NanoFilt v2.8.0 (8) with the parameters “-q 10 --headcrop 50 -l 1000”.

Editor Atmika Paudel, Fluxus Inc., Sunnyvale, California, USA

Address correspondence to Takahiro Kawanabe, kawanabe.takahiro.s@tokai.ac.jp.

Hideki Kinoshita is the manager of Probio Co., Ltd., a startup from Tokai University. Yuki Nakashima holds more than 5% of the shares in the company. The other authors declare no conflicts of interest.

Received 3 July 2025

Accepted 25 October 2025

Published 12 November 2025

Copyright © 2025 Hibi 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 Accession numbers and sequencing statistics

Bacterial species	<i>Pediococcus pentosaceus</i>
Strain	TOKAI 10 m
Assembly statistics	
Total length (bp)	1,879,269
Chromosome length (bp)	1,788,781
Plasmid 1 length (bp)	56,333
Plasmid 2 length (bp)	13,649
Plasmid 3 length (bp)	10,458
Plasmid 4 length (bp)	10,048
No. of contigs	5
Chromosome mean depth (×)	383.353
Plasmid 1 mean depth (×)	250.608
Plasmid 2 mean depth (×)	1,242.73
Plasmid 3 mean depth (×)	477.87
Plasmid 4 mean depth (×)	629.244
GC content (%)	37.2
No. of CDSs	1,918
No. of rRNAs	15
No. of tRNAs	55
No. of CRISPRs	1
Completeness check (%)	91.49
Taxonomy Check (%)	98.7
(ANI to <i>Pediococcus pentosaceus</i> NBRC 107768 [GCA_007992275.1])	

Hybrid genome assembly was performed using Unicycler v0.5.1 (9), and the assembly graph was visualized using Bandage v0.8.1 (10). Trimmed short reads were mapped using BWA-MEM2 v2.2.1 (11), indexed using SAMtools v1.21 (12), and employed for polishing with Pilon v1.24 (13). Genome completeness, taxonomy, and annotation were determined using DFAST v1.3.6 (14), with the chromosome rotated to initiate at the *dnaA* gene. Genome coverage was calculated by mapping trimmed reads using minimap2 v2.29 (15) (16) and SAMtools v1.21 (12). Default parameters were used for all software unless otherwise noted. Annotated genomic features are summarized in Table 1.

Finally, virulence and antibiotic resistance genes were screened using PathogenFinder v1.1 (17) and AMRFinderPlus v4.0.23 (database version: 2025-07-16.1) (18) with the parameter “-i 0.6.” However, no virulence or antibiotic resistance genes were detected in this study.

ACKNOWLEDGMENTS

This study was supported by funding from the Research Institute of Agriculture at Tokai University, as well as by separate research funding from Tokai University.

AUTHOR AFFILIATIONS

¹Graduate School of Bioscience, Tokai University, Kumamoto, Japan

²Research Institute of Agriculture, Tokai University, Kumamoto, Japan

³Undergraduate School of Agriculture, Tokai University, Kumamoto, Japan

⁴JSPS Research Fellowship for Young Scientists, Tokai University, Kumamoto, Japan

⁵Probio Co. Ltd, Kumamoto, Japan

AUTHOR ORCIDs

Tomoyuki Hibi  <http://orcid.org/0009-0009-9719-5564>

Hideki Kinoshita  <http://orcid.org/0000-0002-7240-3832>

Takahiro Kawanabe  <http://orcid.org/0009-0006-7245-7802>

DATA AVAILABILITY

Data availability. All datasets data sets have been deposited in DDBJ/ENA/GenBank, and the accession numbers are listed below. BioProject: [PRJDB35427](https://www.ncbi.nlm.nih.gov/bioproject/PRJDB35427) BioSample: [SAMD01038017](https://www.ncbi.nlm.nih.gov/biosample/SAMD01038017) Raw short reads data: [RRR698281](https://www.ncbi.nlm.nih.gov/sra/RRR698281) Raw long-reads data: [RRR698282](https://www.ncbi.nlm.nih.gov/sra/RRR698282) Assembled and annotated reads data: [AP041085](https://www.ncbi.nlm.nih.gov/sra/AP041085), [AP041086](https://www.ncbi.nlm.nih.gov/sra/AP041086), [AP041087](https://www.ncbi.nlm.nih.gov/sra/AP041087), [AP041088](https://www.ncbi.nlm.nih.gov/sra/AP041088), and [AP041089](https://www.ncbi.nlm.nih.gov/sra/AP041089).

REFERENCES

- Kinoshita H, Sohma Y, Ohtake F, Ishida M, Kawai Y, Kitazawa H, Saito T, Kimura K. 2013. Biosorption of heavy metals by lactic acid bacteria and identification of mercury binding protein. *Res Microbiol* 164:701–709. <https://doi.org/10.1016/j.resmic.2013.04.004>
- Nakashima Y, Shiiyama N, Urabe T, Yamashita H, Yasuda S, Igoshi K, Kinoshita H. 2020. Functions of small RNAs in *Lactobacillus casei-Pediococcus* group of lactic acid bacteria using fragment analysis. *FEMS Microbiol Lett* 367:fnaa154. <https://doi.org/10.1093/femsle/fnaa154>
- Reed DL, Hafner MS. 2002. Phylogenetic analysis of bacterial communities associated with ectoparasitic chewing lice of pocket gophers: a culture-independent approach. *Microb Ecol* 44:78–93. <https://doi.org/10.1007/s00248-002-0009-4>
- Yamamoto N, Shoji M, Hoshigami H, Watanabe K, Watanabe K, Takatsuzu T, Yasuda S, Igoshi K, Kinoshita H. 2019. Antioxidant capacity of soymilk yogurt and exopolysaccharides produced by lactic acid bacteria. *Biosci Microbiota Food Health* 38:97–104. <https://doi.org/10.12938/bmfh.18-017>
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- 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>
- 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>
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>
- Vasimuddin M, Misra S, Li H, Aluru S. 2019. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. Rio de Janeiro, Brazil 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS). <https://doi.org/10.1109/IPDPS.2019.00041>
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008. <https://doi.org/10.1093/gigascience/giab008>
- 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>
- Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Li H. 2021. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* 37:4572–4574. <https://doi.org/10.1093/bioinformatics/btab705>
- Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. 2013. PathogenFinder—distinguishing friend from foe using bacterial whole genome sequence data. *PLoS One* 8:e77302. <https://doi.org/10.1371/journal.pone.0077302>
- Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J, Haft DH, Hoffmann M, Pettengill JB, Prasad AB, Tillman GE, Tyson GH, Klimke W. 2021. AMRFinderPlus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci Rep* 11:12728. <https://doi.org/10.1038/s41598-021-91456-0>