

Complete genome sequence of *Pediococcus pentosaceus* 13.7 13A-2 isolated from a Jersey milk cattle

Sebastian W. Fischer,^{1,2} Nadine Mariani Corea,² Jennifer Wachtarczyk,² Fritz Titgemeyer²

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We present the genome sequence of *Pediococcus pentosaceus* 13.7 13A-2, isolated from a Jersey dairy cow from a conventionally managed dairy farm in Münsterland, Germany. The circular chromosome comprises 1.77 Mbp and 1,686 predicted protein-coding genes. It provides the molecular basis for investigating its potential application in foods.

KEYWORDS lactic acid bacteria, probiotics, protective cultures, Nanopore, biofilm

Pediococcus pentosaceus is a lactic acid bacterium used to ferment milk, meat, vegetables, and fruits (1, 2). It shows probiotic activity and improves food safety by inhibiting foodborne pathogens (2–4).

The strain was isolated from a swab introduced into the teat canal to obtain the biofilm of the bovine papillary duct from a cow from a conventionally managed milk farm in Münsterland, Germany (GPS: 51.95N 7.63E). The swab was transferred to 1 mL of 0.9% NaCl solution. Serial dilutions were plated onto De Man, Rogosa, and Sharpe (MRS)-Bouillon-agar plates supplemented with 0.5 g/L cysteine and 10 mg/L bromophenol blue and incubated anaerobically at 30°C for 48 h (5). Colonies were clonally isolated and subjected to colony PCR for amplification of an *rrnA* gene fragment using primers 27f (5'-AGAGTTTGATCCTGGCTCAG-3') and N1492R (5'-TACGGYTACCTTGTTAYGACTT-3') (6). The Sanger sequenced DNA of 1,101 bp from base 77 to 1,177 was 100% identical to *rrnA* of *P. pentosaceus* ATCC 25745 (NC_008525.1).

The strain was cultured on MRS agar with 0.5 g/L cysteine for 48 h at 30°C. Genomic DNA was isolated using the Wizard HMW DNA Extraction Kit (Promega, USA) and quantified on a DeNovix QFX fluorometer (DeNovix, USA). The sequencing library was prepared using the Oxford Nanopore SQK-RAD004 kit. The kit's transposase fragmented the genomic DNA, and no additional size selection was performed. The sample was loaded (30 µL) onto a Flongle Flow Cell R9.4.1 (FLO-FLG001) mounted on a MinION Mk1B device running MinKNOW v22.08.9 (7). Raw signals were base-called with Guppy v6.3.8 (high-accuracy model dna_r9.4.1_450bps_hac.cf g) (8). Reads with a mean Phred quality ≥Q9 were retained. Adapter trimming was performed with Porechop v0.2.4 (9). Reads were further filtered with Filtlong v0.2.1 to discard the worst 10% by quality and all reads <1 kb (10). The final data set comprised 59,877 reads (240 Mbp; read N50 = 5,834 bp; median/mean read length = 4,008/2,569 bp; longest read = 53,152 bp; mean read Q = 13.1; 59.2%/16.5% of bases reached ≥Q20/≥Q30). Quality metrics were obtained with Nanoq v0.10.0 and SeqKit v2.9.0 (11, 12). Assembly was managed with a Snakemake workflow that implemented Tricycler v0.5.4. Sub-assemblies were generated with Flye v2.9.2, Raven v1.8.2, and Miniasm v0.3, followed by MiniPolish v0.1.3; consensus reconciliation used Mash v2.3, minimap2 v2.25, and MUSCLE v3.8 under R v4.3.1 (packages *ape* v5.7-1 and *phangorn* v2.11-1) (13–26). The draft consensus was polished with Medaka v1.8.0 and homopolished v0.4.1 in mode polish and modpolish (27, 28). The circular chromosome was re-oriented to start at *dnaA* using Dnaapler v1.2.0.

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

Address correspondence to Sebastian W. Fischer, sebastian.fischer@fh-muenster.de.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 30 July 2025

Accepted 13 October 2025

Published 5 November 2025

Copyright © 2025 Fischer 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 Assembly and quality metrics for the complete chromosome of *P. pentosaceus* 13.7 13A-2

Replicon	Length (bp)	G+C (%)	BUSCO completeness (%)	CheckM2 completeness/contamination (%)	Genome fraction (%)
Chromosome	1,771,625	37.14	98.80	100/0.23	85.61

The correctness was verified with Bandage v0.9.0. (29, 30). Unless otherwise noted, all bioinformatics tools were used with default settings.

QUAST v5.3.0, using *P. pentosaceus* ATCC 25745 ([NC_008525.1](https://ncbi.nlm.nih.gov/nucl/NC_008525.1)) as reference, yielded the metrics shown in Table 1; BUSCO v5.8.3 (data set *pediococcus_odt12*) reported 98.80% completeness (31, 32). CheckM2 v1.1.0 (specific neural-network model) confirmed 100% completeness and 0.23% contamination (33). Read depth, obtained by remapping filtered reads with minimap2 v2.30 and samtools v1.22.1, showed a mean coverage of 113.9x across the chromosome (21, 22, 34).

Genome features were annotated with PGAP v6.10, identifying 1,686 protein-coding sequences, 15 rRNA genes, and 55 tRNA genes (35).

ACKNOWLEDGMENTS

We are grateful to Sophie Simon, Lennard Stach, and Alexander Probst for introducing us to the secrets of nanopore sequencing. This publication was supported by the Open Access Publication Fund of the University of Bonn.

S.W.F., J.W., and N.M.C.: Data curation, Formal analysis, Investigation, and Methodology. S.W.F.: Genome assembly and quality control. S.W.F., N.M.C., and F.T.: Writing—original draft, review, and editing.

AUTHOR AFFILIATIONS

¹Institute for Hygiene and Public Health, University Clinics Bonn, Bonn, Germany

²Department of Food, Nutrition, Facilities, FH Münster, Münster, Germany

AUTHOR ORCIDs

Sebastian W. Fischer  <http://orcid.org/0000-0002-4450-5487>

Nadine Mariani Corea  <http://orcid.org/0009-0007-6530-0476>

Fritz Titgemeyer  <http://orcid.org/0009-0002-6235-398X>

FUNDING

Funder	Grant(s)	Author(s)
Open Access Publication Fund of the University of Bonn		Sebastian W. Fischer

AUTHOR CONTRIBUTIONS

Sebastian W. Fischer, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review and editing | Nadine Mariani Corea, Data curation, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review and editing | Jennifer Wachtarczyk, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Fritz Titgemeyer, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Whole-genome sequencing data for *Pedococcus pentosaceus* strain 13.7 13A-2 have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number [PRJNA1297324](https://ncbi.nlm.nih.gov/bioproject/PRJNA1297324). The raw sequencing reads are available via the SRA under the corresponding BioSample [AMN50221057](https://ncbi.nlm.nih.gov/biosample/AMN50221057) entry. The assembled and annotated genome is available at DDBJ/ENA/GenBank accession numbers [JBPWPB000000000](https://ncbi.nlm.nih.gov/nucl/JBPWPB000000000),

NZ_CM1266291, and GCF_052402805.1. The version described in this paper is version JBPWPB01000000.

ETHICS APPROVAL

No Institutional Animal Care and Use Committee (IACUC) approval was required for this study.

REFERENCES

- Mgomi FC, Yuan L, Wang Y, Rao S, Yang Z. 2022. Physiological properties, survivability and genomic characteristics of *Pediococcus pentosaceus* for application as a starter culture. *Int J of Dairy Tech* 75:588–602. <https://doi.org/10.1111/1471-0307.12864>
- Qi Y, Huang L, Zeng Y, Li W, Zhou D, Xie J, Xie J, Tu Q, Deng D, Yin J. 2021. *Pediococcus pentosaceus*: screening and application as probiotics in food processing. *Front Microbiol* 12:762467. <https://doi.org/10.3389/fmicb.2021.762467>
- Ahn SB, Jun DW, Kang B-K, Lim JH, Lim S, Chung M-J. 2019. Randomized, double-blind, placebo-controlled study of a multispecies probiotic mixture in nonalcoholic fatty liver disease. *Sci Rep* 9:5688. <https://doi.org/10.1038/s41598-019-42059-3>
- Olaoye OA, Dodd CER. 2010. Evaluation of bacteriocinogenic *Pediococcus acidilactici* as protective culture in the preservation of *Tsire*, a traditional nigerian stick meat. *J Food Saf* 30:867–888. <https://doi.org/10.1111/j.1745-4565.2010.00247.x>
- Lee HM, Lee Y. 2008. A differential medium for lactic acid-producing bacteria in a mixed culture. *Lett Appl Microbiol* 46:676–681. <https://doi.org/10.1111/j.1472-765X.2008.02371.x>
- Mariani Coreia N, Titgemeyer F, Wachtarczyk J, Meyer F, Fischer SW. 2025. Complete genome sequence of *Pediococcus pentosaceus* 13.7 2A-1 isolated from a Holstein Friesian dairy cattle. *Microbiol Resour Announc* e0084825. <https://doi.org/10.1128/mra.00848-25>
- Oxford Nanopore Technologies. 2022. MinKNOW (version 22.08.9). Oxford Nanopore Technologies. Available from: <https://nanoporetech.com/software>
- Oxford Nanopore Technologies. 2022. Guppy (version 6.3.8). Oxford Nanopore Technologies. Available from: <https://nanoporetech.com/software/other/guppy>
- Wick RR. 2018. Porechop (version 0.2.4). GitHub. Available from: <https://github.com/rrwick/Porechop>
- Wick RR. 2021. Filtlong (version 0.2.1). GitHub. Available from: <https://github.com/rrwick/Filtlong>
- Shen W, Sipos B, Zhao L. 2024. SeqKit2: a Swiss army knife for sequence and alignment processing. *Imeta* 3:e191. <https://doi.org/10.1002/imt.2191>
- Steinig E, Coin L. 2022. Nanoq: ultra-fast quality control for nanopore reads. *JOSS* 7:2991. <https://doi.org/10.21105/joss.02991>
- Mölder F, Jablonski KP, Letcher B, Hall MB, van Dyken PC, Tomkins-Tinch CH, Sochat V, Forster J, Vieira FG, Meesters C, Lee S, Twardziok SO, Kanitz A, VanCampen J, Malladi V, Wilm A, Holtgrewe M, Rahmann S, Nahnsen S, Köster J. 2021. Sustainable data analysis with Snakemake. *F1000Res* 10:33. <https://doi.org/10.12688/f1000research.29032.3>
- Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 22:266. <https://doi.org/10.1186/s13059-021-02483-z>
- 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>
- Vaser R, Šikić M. 2021. Time- and memory-efficient genome assembly with Raven. *Nat Comput Sci* 1:332–336. <https://doi.org/10.1038/s43588-021-00073-4>
- Li H. 2016. Minimap and miniasm: fast mapping and *de novo* assembly for noisy long sequences. *Bioinformatics* 32:2103–2110. <https://doi.org/10.1093/bioinformatics/btw152>
- Wick RR, Holt KE. 2021. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. *F1000Res* 8:2138. <https://doi.org/10.12688/f1000research.21782.4>
- Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 17:132. <https://doi.org/10.1186/s13059-016-0997-x>
- Ondov BD, Starrett GJ, Sappington A, Kostic A, Koren S, Buck CB, Phillippy AM. 2019. Mash Screen: high-throughput sequence containment estimation for genome discovery. *Genome Biol* 20:232. <https://doi.org/10.1186/s13059-019-1841-x>
- 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>
- Edgar RC. 2022. High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *bioRxiv*. <https://doi.org/10.1101/2021.06.20.449169>
- R Core Team. 2021. R: a language and environment for statistical computing (4.1.2). Vienna, Austria. R Foundation for Statistical Computing.
- Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35:526–528. <https://doi.org/10.1093/bioinformatics/bty633>
- Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27:592–593. <https://doi.org/10.1093/bioinformatics/btq706>
- Oxford Nanopore Technologies. 2024. Medaka (versions 1.8 and 2.1). Oxford Nanopore Technologies. <https://github.com/nanoporetech/medaka>
- Huang Y-T, Liu P-Y, Shih P-W. 2021. Homopolish: a method for the removal of systematic errors in nanopore sequencing by homologous polishing. *Genome Biol* 22:95. <https://doi.org/10.1186/s13059-021-02282-6>
- Bouras G, Grigson SR, Papudeshi B, Mallawaarachchi V, Roach MJ. 2024. Dnaapler: a tool to reorient circular microbial genomes. *JOSS* 9:5968. <https://doi.org/10.21105/joss.05968>
- 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>
- 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>
- Manni M, Berkeley MR, Seppely M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *arXiv*. <https://doi.org/10.48550/arXiv.2106.11799>
- Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 20:1203–1212. <https://doi.org/10.1038/s41592-023-01940-w>
- 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>
- Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretin 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>