

# Complete genome sequence of *Citroniella saccharovorans* DSM 29873, isolated from human fecal sample

George Cheng,<sup>1</sup> Maria Westerholm,<sup>1</sup> Anna Schnürer<sup>1</sup>

**AUTHOR AFFILIATION** See affiliation list on p. 2.

**ABSTRACT** A complete genome was recovered from *Citroniella saccharovorans*, strain DSM 29873, using Oxford Nanopore Technologies. The genome assembly contains 1,413,868 bp with 30.23% G+C content. The species belongs to the family *Peptoniphilaceae* and, as of yet, is the only cultivated representative of the genus *Citroniella*.

**KEYWORDS** human microbiome, fecal organisms, *Citroniella saccharovorans*, Nanopore

*Citroniella saccharovorans* (DSM 29873) was isolated from a fecal sample from a traditional native on the southern coast of Peru (1). *C. saccharovorans* grows under strict anaerobic conditions at optimum 30°C and pH 6.5–7.3, utilizing different sugars, organic acids, and amino acids for growth, with production of acetate and methyl succinate (1). This complete genome will provide possibilities to recover information on genes and potential metabolic pathways used by this species, representing the only isolate of the genus *Citroniella*. This strain was initially enriched anaerobically by inoculation of a fecal sample in liquid medium 2 (1). For isolation, aliquots of the enrichment were transferred to solid agar plates and colonies were finally picked and sub-cultured on the same medium. Patel et. al. (1) performed phylogenetic analysis of the 16S rRNA gene (NR\_173562) using UltraClean Microbial DNA Isolation Kit (MoBio Laboratories) and PCR amplification using the universal primers 8F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTACCTTGTTACGACTT-3') (1).

*C. saccharovorans* was retrieved from DSMZ (DSM 29873) as a lyophilized strain and reconstituted following DSMZ instructions, then streaked onto Columbia Blood Agar (ThermoFisher Scientific Columbia Blood Agar with sheep blood medium, PB5039A) in an anaerobic chamber (Whitley A45 anaerobic workstation) at 37°C. After 5 days, a single colony was used for colony PCR amplification (PuReTaq Read-To-Go PCR beads, Cytiva, UK) using primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R. The PCR cycle was initiated with denaturing step at 95°C for 15 min, 30 cycles at 95°C, then 55°C and 72°C for 1 min each step, finishing with an extension step at 72°C for 7 min (2). Sanger sequencing of the PCR product resulted in a 1,400 bp fragment, showing 99.42% identity to *C. saccharovorans* strain DSM 29873.

DNA was recovered using the NucleoBond kit (Macharey Nagal) following extraction protocol from Sun et al. (3), with modifications in steps 1–6, replaced with resuspension of colonies from solid medium into 2.5 mL Eppendorf tube with Buffer G3. Oxford Nanopore MinION ligation sequencing kit (SQK-LSK109) and native barcoding expansion (EXP-NBD104) were used for sequencing without shearing and size selection. Flow cell FLO-MIN106D R9v and the software was running MinKNOW (v23.04.6) with GUPPY (v6.5.7) basecalling with High Accuracy.

The genome assembly was completed by Nanozoo GmbH. All tools were used with default parameters unless stated otherwise. Fastq\_utils (v0.24.1) (4) and seqkit (v0.13.2) (5) were used for validation and manipulation of fastq files, filtlong (v0.2.0) (6) was used to filter long reads by quality, bmap (v38.86) (7) was used for read alignment,

**Editor** Vanja Klepac-Ceraj, Wellesley College, Wellesley, Massachusetts, USA

Address correspondence to Anna Schnürer, anna.schnurer@slu.se.

The authors declare no conflict of interest.

See the funding table on p. 2.

**Received** 14 January 2024

**Accepted** 20 February 2024

**Published** 11 March 2024

Copyright © 2024 Cheng 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/).

and Flye (v2.9.1) (8) was used for genome assembly and created consensus sequences with medaka (v1.9.1) (9). Annotation was completed with NCBI prokaryotic genome annotation pipeline (PGAP) (10). Bandage (v0.8.1) (11) was used for visualizing *de novo* assembly graphs. The completeness of the genome was verified with CheckM (v1.1.3) (12). The assembled genome size is 1,953,712 bp with a genome coverage of 40× and a completeness of 98.42% and 0% contamination. The final assembly contains three contigs with GC content of 30.23% and N<sub>50</sub> of 1,413,868 bp. The annotated genome has 1,939 predicted genes, 1,893 coding sequences (CDSs), 37 tRNA sequences, two 5S, two 16S, and two 23S rRNA genes.

# ACKNOWLEDGMENTS

The authors acknowledge Christian Brandt from Nanozoo GmbH for the reconstruction of the genome.

This work was supported and funded by Formas (2018–01341) (to G.C. and A.S.) and European Research Council under grant number 948138 (to M.W.) and by the Swedish University of Agricultural Sciences (SLU).

# AUTHOR AFFILIATION

<sup>1</sup>Department of Molecular Sciences, Biocenter Swedish University of Agricultural Sciences, Uppsala, Sweden

# AUTHOR ORCIDs

George Cheng  <http://orcid.org/0000-0001-8013-8402>  
 Anna Schnürer  <http://orcid.org/0000-0003-0038-553X>

# FUNDING

| Funder  | Grant(s)   | Author(s)        |
|---|------------|------------------|
| <a href="#">Svenska Forskningsrådet Formas (Formas)</a> | 2018-01341 | George Cheng     |
| <a href="#">EC   European Research Council (ERC)</a>    | 948138     | Maria Westerholm |
| <a href="#">Svenska Forskningsrådet Formas (Formas)</a> | 2018-01341 | Anna Schnürer    |

# AUTHOR CONTRIBUTIONS

George Cheng, Data curation, Formal analysis, Writing – original draft | Maria Westerholm, Writing – review and editing | Anna Schnürer, Project administration, Supervision, Writing – review and editing

# DATA AVAILABILITY

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession [JAYKOT000000000](#). The version described in this paper is the first version, [JAYKOT000000000.1](#). The Nanopore reads can be found under BioProject accession [PRJNA1050169](#).

# REFERENCES

- Patel NB, Obregón-Tito AJ, Tito RY, Trujillo-Villaroel O, Marin-Reyes L, Troncoso-Corzo L, Guija-Poma E, Lewis CM, Lawson PA. 2019. *Citroniella saccharovorans* gen. nov. sp. nov., a member of the family *Peptoniphilaceae* isolated from a human fecal sample from a coastal traditional community member. *Int J Syst Evol Microbiol* 69:1142–1148. <https://doi.org/10.1099/ijsem.0.003287>
- Sun H, Bjerketorp J, Levenfors JJ, Schnürer A. 2020. Isolation of antibiotic-resistant bacteria in biogas digestate and their susceptibility to antibiotics. *Environ Pollut* 266:115265. <https://doi.org/10.1016/j.envpol.2020.115265>
- Sun H, Brandt C, Schnürer A. 2020. Long-read DNA preparation for bacterial isolates. <https://doi.org/10.17504/protocols.io.64ghgtw>
- Fonseca N. 2023. Nunofonseca/Fastq\_Utils. C. [https://github.com/nunofonseca/fastq\\_utils](https://github.com/nunofonseca/fastq_utils).
- Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11:e0163962. <https://doi.org/10.1371/journal.pone.0163962>
- Wick R. 2024. "Filtlong: Quality filtering tool for long reads." C++. <https://github.com/rrwick/Filtlong>.

7. Bushnell B. 2014. BBLMap: a fast, accurate, splice-aware aligner. LBNL-7065E. Berkeley, CA (United States) Lawrence Berkeley National Lab. (LBNL)
8. 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>
9. Github - Nanoporetech/Medaka: Sequence correction provided by ONT research. n.d. <https://github.com/nanoporetech/medaka>.
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>
11. 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>
12. 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>