



Complete Genome Sequence of the Industrial Fast-Acidifying Strain *Streptococcus thermophilus* N4L

Lucas Proust,^{a,c} Valentin Loux,^b Véronique Martin,^b Cristian Magnabosco,^d Martin Pedersen,^d Véronique Monnet,^a Vincent Juillard^a

^aMicalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

^bMalAGE, INRA, Université Paris-Saclay, Jouy-en-Josas, France

^cProcelys, Lesaffre Group, Maisons-Alfort, France

^dSacco S.r.l., Cadorego, Italy

ABSTRACT *Streptococcus thermophilus* is one of the most used dairy starters for the production of yogurt and cheese. We report here the complete genome sequence of the industrial strain *S. thermophilus* N4L, which is used in dairy technology for its fast-acidifying phenotype.

Streptococcus thermophilus is a well-known lactic acid bacterium (LAB) widely used for the production of dairy products in association with other LAB species, such as *Lactococcus lactis* and *Lactobacillus delbrueckii* subsp. *bulgaricus*, for cheese and yogurt production, respectively. It belongs to the salivarius group and recently evolved from a commensal ancestor (1). Nowadays, it can be naturally found all around the world in dairy production facilities, whether artisanal or industrial (2). For industrial use, strains are carefully selected for their ability to ferment milk and confer desirable properties to the final product. In particular, due to the high prevalence of the cell wall proteinase *prtS* gene (1) in industrial isolates of *S. thermophilus*, they are often associated with higher productivity in milk when compared to their wild or traditional counterparts. The proteinase PrtS is responsible for the breakdown of caseins into peptides, which are used as sources of amino acids, thus allowing faster growth and acidification than do PrtS-negative strains. These released peptides are then internalized inside the bacterial cell primarily by the Ami oligopeptide transport system, which includes one or several peptide-binding proteins depending on the strain (3).

S. thermophilus N4L is an industrial strain used precisely for its fast-acidifying phenotype. A stationary-phase culture (M17 with 2% lactose) was subjected to genomic DNA (gDNA) extraction using the GenElute bacterial gDNA kit (Sigma-Aldrich, USA) following the protocol indicated by the manufacturer. Its gDNA was then sequenced (Macrogen, Inc., South Korea) on a PacBio RS II system (Pacific Biosciences, USA) using single-molecule real-time (SMRT) sequencing technology (4). A total of 150,292 polymerase reads were obtained, displaying an average length of 11,243 bp. These reads, after quality filtering and adapter trimming, were converted into 213,700 subreads with an average length of 7,626 bp. They were subsequently used for *de novo* genome assembly. For that purpose, Canu version 1.6 was employed with all default settings for subread correction, trimming, and assembly (5). The resulting genome corresponded to a unique circular chromosome of 1,831,756 bp with an 889× coverage depth and a G+C content of 39.05%.

Annotation was performed with the AGMIAL annotation platform (6). It revealed the presence of 2,300 open reading frames (ORFs), including 67 tRNAs and 18 rRNAs. This sequencing work also confirmed the presence of the genomic island encoding the *prtS* gene, as well as two copies of the oligopeptide-binding proteins AmiA1 and AmiA3, which explains the high acidification rates in milk observed with this strain. All of the

Received 20 July 2018 Accepted 27 July 2018 Published 30 August 2018

Citation Proust L, Loux V, Martin V, Magnabosco C, Pedersen M, Monnet V, Juillard V. 2018. Complete genome sequence of the industrial fast-acidifying strain *Streptococcus thermophilus* N4L. Microbiol Resour Announc 7:e01029-18. <https://doi.org/10.1128/MRA.01029-18>.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Copyright © 2018 Proust 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/).

Address correspondence to Vincent Juillard, vincent.juillard@inra.fr.

genes required for natural competence are also present in the genome (7, 8), suggesting that the strain could theoretically be naturally transformable.

Data availability. The complete genome sequence of *S. thermophilus* strain N4L has been deposited in GenBank under the accession no. [LS974444](https://www.ncbi.nlm.nih.gov/GenBank/LS974444) (BioProject no. [PRJEB27286](https://www.ncbi.nlm.nih.gov/BioProject/PRJEB27286)).

ACKNOWLEDGMENTS

The bacterial strain was provided by Sacco S.r.l., which also supported the sequencing costs. This work was performed in the framework of a project funded by the Association Nationale de la Recherche et de la Technologie (ANRT).

REFERENCES

1. Delorme C, Bartholini C, Bolotine A, Ehrlich SD, Renault P. 2010. Emergence of a cell wall protease in the *Streptococcus thermophilus* population. *Appl Environ Microbiol* 76:451–460. <https://doi.org/10.1128/AEM.01018-09>.
2. Delorme C, Legravet N, Jamet E, Hoarau C, Alexandre B, El-Sharoud WM, Darwish MS, Renault P. 2017. Study of *Streptococcus thermophilus* population on a world-wide and historical collection by a new MLST scheme. *Int J Food Microbiol* 242:70–81. <https://doi.org/10.1016/j.ijfoodmicro.2016.11.016>.
3. Garault P, Le Bars D, Besset C, Monnet V. 2002. Three oligopeptide-binding proteins are involved in the oligopeptide transport of *Streptococcus thermophilus*. *J Biol Chem* 277:32–39. <https://doi.org/10.1074/jbc.M107002200>.
4. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* 323:133–138. <https://doi.org/10.1126/science.1162986>.
5. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
6. Bryson K, Loux V, Bossy R, Nicolas P, Chaillou S, van de Guchte M, Penaud S, Maguin E, Hoebeke M, Bessières P, Gibrat J-F. 2006. AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system. *Nucleic Acids Res* 34:3533–3545. <https://doi.org/10.1093/nar/gkl471>.
7. Fontaine L, Wahl A, Flécharde M, Mignolet J, Hols P. 2015. Regulation of competence for natural transformation in streptococci. *Infect Genet Evol* 33:343–360. <https://doi.org/10.1016/j.meegid.2014.09.010>.
8. Gardan R, Besset C, Gitton C, Guillot A, Fontaine L, Hols P, Monnet V. 2013. Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. *J Bacteriol* 195:1845–1855. <https://doi.org/10.1128/JB.02196-12>.