

Genome Sequence of *Lactobacillus fermentum* Strain NCC2970 (CNCM I-5068)

Caroline Barretto,^a Catherine Ngom-Bru,^a Alienor Genevaz,^a Coralie Fournier,^b Deborah Moine,^b Mohamed Kassam,^b Agnes Ittis,^c Pierre Sagory-Zalkind,^c Pierre-Emmanuel Ciron,^c Gilles Faucherand,^c Patrick Descombes,^b Stephane Duboux^a

Nestec Ltd., Nestlé Research Center, Lausanne, Switzerland^a; Nestlé Institute of Health Sciences, Lausanne, Switzerland^b; Genostar, Montbonnot-Saint-Martin, France^c

***Lactobacillus fermentum* NCC2970 (CNCM I-5068) is a lactic acid bacterium originating from the Nestlé Culture Collection. Here, we disclose its full 1.9-Gb genome sequence comprising one chromosome with no plasmid.**

Received 14 September 2016 Accepted 26 September 2016 Published 17 November 2016

Citation Barretto C, Ngom-Bru C, Genevaz A, Fournier C, Moine D, Kassam M, Ittis A, Sagory-Zalkind P, Ciron P-E, Faucherand G, Descombes P, Duboux S. 2016. Genome sequence of *Lactobacillus fermentum* strain NCC2970 (CNCM I-5068). *Genome Announc* 4(6):e01254-16. doi:10.1128/genomeA.01254-16.

Copyright © 2016 Barretto 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 Caroline Barretto, caroline.barretto@rdls.nestle.com.

Lactobacillus fermentum is a heterofermentative lactic acid bacterium (LAB) belonging to the *Lactobacillaceae* family within the *Bacilli* class of the *Firmicutes* phylum. The species is listed in the qualified presumption of safety (QPS) published by the European Food Safety Authority (EFSA) (1), and a previously described strain belonging to this species has been generally recognized as safe by the U.S Food and Drug Administration (GRAS notice no. 531). *L. fermentum* strains have been previously reported for their different properties, including technological (2) or probiotic functionalities (3–6). *Lactobacillus fermentum* NCC2970 originates from the Nestlé Culture Collection and has also been deposited at the National Collection of Microorganisms Cultures (CNCM) under CNCM I-5068.

Genomic DNA was extracted from mid-exponential cultures using a Gentra DNA Purgene kit (Qiagen), and 20-kb libraries were prepared following the Pacific Biosciences (PacBio) protocol and BluePippin size selection (Sage Science). Sequencing was performed on the PacBio RSII platform using P6/C4 chemistry on single-molecule real-time (SMRT) cells with a 240-min collection protocol. The subreads were *de-novo* assembled using the PacBio Hierarchical Genome Assembly Process (HGAP)/Quiver software package (7) followed by Circlator for genome circularization (8, 9) and a final polishing step with Quiver. The strain was assembled into a single contig corresponding to the chromosome. No plasmid could be detected.

During sequencing, epigenetic modifications of each nucleotide position were measured as kinetic variations (KVs) in nucleotide incorporation rates. Motifs were deduced from the KV data (10). Analysis was performed using the SMRT portal RS_Modification_and_Motif_Analysis protocol (PacBio).

The length of the chromosome of *L. fermentum* NCC2970 (CNCM I-5068) is 1,949,874 bp, with a GC content of 52.21%. Automated expert annotation carried out by Genostar using a proprietary pipeline (2) revealed 5 rRNA operons and 56 tRNA genes. It indicated 1,927 protein-coding sequences on the chromosome, of which 1,650 (86%) were annotated with known biological functions and 277 (14%) encode hypothetical proteins or uncharacterized proteins.

The *L. fermentum* NCC2970 (CNCM I-5068) genome has a similar size and structure than other publically available *L. fermentum* genomes (F-6 [5], CECT 5716 [6]). It contains, however, some specific features, especially with regard to its substrate usage capacity. Compared to other *L. fermentum* genomes, it harbors a more complete shikimate pathway with a 3-dehydroquinate dehydratase (EC 4.2.1.10; CDS0787) and a shikimate dehydrogenase (EC 1.1.1.25; CDS0788), which are, respectively, the third and the fourth enzymes of this pathway, enabling the conversion of quinic acid to shikimic acid. It harbors also a 5'-nucleotidase (EC 3.1.3.5; CDS0645), which in *C. glutamicum* is required for growth when nucleotides are provided as the sole source of phosphate (11), as well as a phosphopentomutase (EC 5.4.2.7; CDS0031), which has been shown in *Bacillus cereus* to be necessary to isomerize ribose-1P in ribose-5P, thus enabling it to enter the pentose phosphate pathway (12).

Accession number(s). This whole-genome project has been deposited at DDBJ/EMBL/GenBank under the accession number CP017151. The version described in this paper is the first version.

REFERENCES

- EFSA Panel on Biological Hazards (BIOHAZ). 2015. Statement on the update of the list of QPS-recommended biological agents intentionally added to food or feed as notified to EFSA 3: suitability of taxonomic units notified to EFSA until September 2015. *EFSA J* 13:4331. <http://dx.doi.org/10.2903/j.efsa.2015.4331>.
- Owusu-Kwarteng J, Tano-Debrah K, Akabanda F, Jespersen L. 2015. Technological properties and probiotic potential of *Lactobacillus fermentum* strains isolated from West African fermented millet dough. *BMC Microbiol* 15: <http://dx.doi.org/10.1186/s12866-015-0602-6>.
- Lin WH, Yu B, Jang SH, Tsen HY. 2007. Different probiotic properties for *Lactobacillus fermentum* strains isolated from swine and poultry. *Anaerobe* 13:107–113. <http://dx.doi.org/10.1016/j.anaerobe.2007.04.006>.
- Lopez-Huertas E. 2015. Safety and efficacy of human breast milk *Lactobacillus fermentum* CECT 5716: a mini-review of studies with infant formulae. *Benef Microbes* 6:219–224. <http://dx.doi.org/10.3920/BM2014.0091>.
- Sun Z, Zhang W, Bilige M, Zhang H. 2015. Complete genome sequence of the probiotic *Lactobacillus fermentum* F-6 isolated from raw milk. *J Biotechnol* 194:110–111. <http://dx.doi.org/10.1016/j.jbiotec.2014.12.010>.
- Jiménez E, Langa S, Martín V, Arroyo R, Martín R, Fernández L, Rodríguez JM. 2010. Complete genome sequence of *Lactobacillus fermentum*

- tum* CECT 5716, a probiotic strain isolated from human milk. *J Bacteriol* 192:4800. <http://dx.doi.org/10.1128/JB.00702-10>.
7. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <http://dx.doi.org/10.1038/nmeth.2474>.
 8. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <http://dx.doi.org/10.1186/s13059-015-0849-0>.
 9. Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics* 8:64. <http://dx.doi.org/10.1186/1471-2105-8-64>.
 10. Korlach J, Turner SW. 2012. Going beyond five bases in DNA sequencing. *Curr Opin Struct Biol* 22:251–261. <http://dx.doi.org/10.1016/j.sbi.2012.04.002>.
 11. Bogan KL, Brenner C. 2010. 5'-nucleotidases and their new roles in NAD⁺ and phosphate metabolism. *New J Chem* 34:845–853. <http://dx.doi.org/10.1039/B9NJ00758J>.
 12. Ipata PL, Sgarrella F, Catalani R, Tozzi MG. 1983. Induction of phosphoribomutase in *Bacillus cereus* growing on nucleosides. *Biochim Biophys Acta* 755:253–256. [http://dx.doi.org/10.1016/0304-4165\(83\)90211-8](http://dx.doi.org/10.1016/0304-4165(83)90211-8).