Draft Genome Sequence of *Lactobacillus casei* DPC6800, an Isolate with the Potential to Diversify Flavor in Cheese

Ewelina Stefanovic,[a,b] Aidan Casey,[a] Paul Cotter,[a] Daniel Cavanagh,[a] Gerald Fitzgerald,[b] Olivia McAuliffe[a]

Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland[a]; School of Microbiology, University College Cork, Cork, Ireland[b]

***Lactobacillus casei* is a nonstarter lactic acid bacterium commonly present in various types of cheeses. It is believed that strains of this species have a significant impact on the development of cheese flavor. The draft genome sequence of *L. casei* DPC6800, isolated from a semi-hard Dutch cheese, is reported.**

Address correspondence to Olivia McAuliffe, olivia.mcauliffe@teagasc.ie.

*L*actobacillus casei is a member of the lactic acid bacteria, a group of Gram-positive, facultatively anaerobic and fastidious bacteria with many biotechnological and health-related applications (1). Strains of the *L. casei* species show extraordinary niche adaptability and have been found in various habitats, such as milk and dairy products, plant materials, and in the human and animal gastrointestinal tracts (1, 2). In dairy products, this organism forms part of the nonstarter microbial flora, which has a prominent role during cheese ripening in the development of specific flavor and aroma compounds (3) through the breakdown of numerous substrates, such as amino acids, fatty acids, and carbohydrates, during cheese production and ripening (4). The subject of this analysis, *L. casei* DPC6800, was isolated from a semi-hard Dutch cheese.

Bacterial DNA from strain DPC6800 was extracted, and single-end sequencing was performed on a Roche 454 FLX sequencer housed in the Teagasc Sequencing Centre (Moorepark, Fermoy, Cork, Ireland) using standard protocols from the manufacturer (Roche Diagnostics Ltd., West Sussex, United Kingdom). Quality filtering, adapter clipping, and trimming of the resulting reads as well as assembly were performed using the SeqMan NGen application of the DNAStar Lasergene Genomics Suite (DNASTAR Inc., Madison). Open reading frames (ORFs) were predicted using Glimmer v3.02 (5) and RAST (6). The genome was annotated using the RAST server, with subsequent annotations verified and manually curated using BLASTp (7) and Artemis (8).

Sequence assembly yielded a 3,053,365 bp draft genome with $31\times$ average coverage, consisting of 58 nonoverlapping contigs with a contig N_{50} of 98,006 bp and a maximum contig size of 595,092 bp. Whole-genome annotation determined that strain DPC6800 contained a total of 3,300 protein-coding genes and 14 tRNAs. Genes that encode enzymes of crucial importance for flavor development were identified, including components of the proteolytic system such as proteinases, peptidases, and aminotransferases. The cell-wall associated proteinase PrtP (AC564_0739c) was identified, along with numerous peptidases of broad or specific peptidolytic function, such as tripeptide aminopepti-dase (AC564_0751c), methionine aminopeptidase (AC564_0890), aminopeptidase S (AC564_0896), aminopeptidase N (AC564_1879c), aminopeptidase V (AC564_3148c), aminopeptidase C (AC564_3291, AC564_3292), and Xaa–Pro-dipeptidyl peptidase (AC564_2630, AC564_2631). Aminotransferases, responsible for the interconversion of amino acids in the later steps of the proteolytic process, are encoded by several genes, i.e., three aspartate aminotransferases (AC564_0742c, AC564_2175, AC564_2467c), two aromatic amino acid aminotransferases (AC564_1682c, AC564_3204), and one branched-chain amino acid aminotransferase (AC564_2001). A gene for glutamate dehydrogenase (AC564_0811c), an enzyme that supports aminotransferase activity through recycling of α-ketoglutarate, an intermediate molecule in aminotransferase reactions, was also identified. Also important for flavor development is the metabolism of citrate, and the presence of a gene encoding an Mg2+-citrate co-transporter CitMHS, necessary for the initial steps of citrate metabolism, was confirmed (AC564_1305). The findings of the genome analysis confirm the potential of *L. casei* DPC6800 for use as an adjunct culture in cheese production to direct or enhance cheese flavor.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number LNQD00000000. The version described in this paper is version LNQD01000000.

REFERENCES

1. **Broadbent JR, Neeno-Eckwall EC, Stahl B, Tandee K, Cai H, Morovic W, Horvath P, Heidenreich J, Perna NT, Barrangou R, Steele JL.** 2012. Analysis of the *Lactobacillus casei* supragenome and its influence in species evolution and lifestyle adaptation. BMC Genomics 13:533. http://dx.doi.org/10.1186/1471-2164-13-533.

2. **Yu S, Peng Y, Zheng Y, Chen W**. 2015. Comparative genome analysis of *Lactobacillus casei*: insights into genomic diversification for niche expansion. Indian J Microbiol **55**:102–107. http://dx.doi.org/10.1007/s12088-014-0496-2.

3. **Settanni L, Moschetti G**. 2010. Non-starter lactic acid bacteria used to improve cheese quality and provide health benefits. Food Microbiol **27**:691–697. http://dx.doi.org/10.1016/j.fm.2010.05.023.

4. **Marilley L, Casey MG**. 2004. Flavours of cheese products: metabolic pathways, analytical tools and identification of producing strains. Int J Food Microbiol **90**:139–159. http://dx.doi.org/10.1016/S0168-1605(03)00304-0.

5. **Delcher AL, Bratke KA, Powers EC, Salzberg SL**. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinformatics **23**:673–679.

6. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD**. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9**:. http://dx.doi.org/10.1186/1471-2164-9-75.

7. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. 1990. Basic local alignment search tool. J Mol Biol **215**:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.

8. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, Barrell B**. 2000. Artemis: sequence visualization and annotation. Bioinformatics **16**:944–945.