Draft Genome Sequence of *Lactobacillus helveticus* Lh 23, Isolated from Natural Whey Starter

Gaia Bertani,[a] Daniela Bassi,[b,c] Claudia Cortimiglia,[b,c] Monica Gatti,[a] Pier Sandro Cocconcelli,[b,c] Erasmo Neviani[a]

aDepartment of Food and Drug, University of Parma, Parma, Italy
bDISTAS, Università Cattolica del Sacro Cuore, Piacenza, Italy
cBiotechnology Research Centre, Cremona, Italy

ABSTRACT *Lactobacillus helveticus* is a thermophilic lactic acid bacterium that is widely employed as a starter culture for manufacturing several Swiss and Italian hard-cooked cheeses. The sequencing of *L. helveticus* Lh 23, which consists of 2,100,230 bp with a GC content of 36.5%, reveals industrially useful traits and interesting metabolic pathways.

*L*actobacillus helveticus is a homofermentative, thermophilic lactic acid bacterium that plays a role as a natural dairy starter in the production of several Swiss and Italian hard-cooked cheeses due to different technological traits encoded in its chromosome (1, 2). The main technological functions of *L. helveticus* in the natural whey starter (NWS) culture environment are lactose depletion during curd acidification and, later, proteolytic activity over the long ripening period. We report the genome sequencing of *L. helveticus* Lh 23, a strain that was isolated from NWS used in the production of Grana Padano, a protected denomination of origin (PDO) Italian cheese (3). This strain, which was previously considered in a comparative study using a microarray of 38 different strains, was genetically different from the other analyzed strains with respect to genes involved in the proteolytic system and in amino acid metabolism, highlighting a high level of genotypic biodiversity due to different environmental and technological selective pressures (3–6). Therefore, a deeper knowledge of its genome may reveal industrially useful traits and interesting metabolism. *Lactobacillus helveticus* Lh 23 was isolated by inoculating 10-fold serial dilutions of NWS onto MRS agar (Oxoid, Waltham, MA, USA) and incubating it anaerobically at 42°C for 48 h. A single colony was picked and used for bacterial DNA extraction, which was performed using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. DNA integrity was checked by agarose gel electrophoresis, and the concentration and purity were determined with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

The sample library was prepared using the Illumina TruSeq DNA sample preparation kit and was sequenced using one lane of a single flow cell on an Illumina HiSeq 1000 system with a 2 × 100-bp paired-end protocol (Genomics Platform; Parco Tecnologico Padano, Lodi, Italy), yielding 11,924,037 reads with an average insert size of 488 bp and 993.0-fold coverage. The genome was directly assembled into 312 contigs with SPAdes software v. 3.1.0; it consisted of 2,100,230 bp with a GC content of 36.5%. The draft genome contained 2,285 genes, including 1,811 protein-coding genes, 403 pseudogenes, 6 rRNAs, 62 tRNAs, and 3 other RNAs (7, 8). Gene functional annotation was performed using BLASTp with the KEGG, Swiss-Prot, and RAST v. 2.0 databases (9–11). The genome analysis revealed several genes involved in amino acid metabolism and in the proteolytic system, i.e., 128 and 39 genes, respectively. RAST annotation showed 187 genes involved in sugar metabolism, transport, and uptake, such as genes partic-

ipating in lactose, galactose, maltose, fructose, chitin, and *N*-acetylglucosamine utilization. Four intact and two incomplete phage regions were identified by employing PHAge Search Tool (PHAST) (12). KEGG analysis revealed that the Lh 23 genome included 7 genes involved in the complete metabolic pathway for the biosynthesis of folate. Moreover, 2 genes involved in exopolysaccharide biosynthesis were identified through RAST annotation. Default parameters were used for all software unless otherwise specified.

Data availability. Data are available from the NCBI under BioProject number PRJNA310689. This whole-genome shotgun project was deposited in DDBJ/ENA/GenBank under accession number LSVJ00000000 and SRA accession number SRR12399863.

REFERENCES

1. Gatti M, Trivisano C, Fabrizi E, Neviani E, Gardini F. 2004. Biodiversity within *Lactobacillus helveticus* isolated from different natural whey starter cultures as revealed by classification trees. Appl Environ Microbiol 70:182–190. https://doi.org/10.1128/aem.70.1.182-190.2004.
2. Giraffa G. 2014. *Lactobacillus helveticus*: importance in food and health. Front Microbiol 5:338. https://doi.org/10.3389/fmicb.2014.00338.
3. Gatti M, Lazzi C, Rossetti L, Mucchetti G, Neviani E. 2003. Biodiversity in *Lactobacillus helveticus* strains in natural whey starter used for Parmigiano Reggiano cheese. J Appl Microbiol 95:463–480. https://doi.org/10.1046/j.1365-2672.2003.01997.x.
4. Lombardi A, Dal Maistro L, De Dea P, Gatti M, Giraffa G, Neviani E. 2002. A poliphasic approach to highlight genotypic and phenotypic diversities of *Lactobacillus helveticus* strains isolated from dairy starter cultures and cheeses. J Dairy Res 69:139–149. https://doi.org/10.1017/s0022029901005349.
5. Broadbent JR, Cai H, Larsen RL, Hughes JE, Welker DL, De Carvalho VG, Tompkins TA, Ardö Y, Vogensen F, De Lorentiis A, Gatti M, Neviani E, Steele JL. 2011. Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains. J Dairy Sci 94:4313–4328. https://doi.org/10.3168/jds.2010-4068.
6. Bertani G, Bassi D, Gatti M, Cocconcelli PS, Neviani E. 2018. Draft genome sequence of *Lactobacillus helveticus* strain Lh 12 isolated from natural whey starter. Genome Announc 6:e00139-18. https://doi.org/10.1128/genomeA.00139-18.
7. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.
8. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–W57. https://doi.org/10.1093/nar/gkw413.
9. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27.
10. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.
11. 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, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
12. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. Nucleic Acids Res 39(Suppl 2):W347–W352. https://doi.org/10.1093/nar/gkr485.