JOURNAL OF BACTERIOLOGY, July 2011, p. 3419–3420 Vol. 193, No. 13
0021-9193/11/$12.00 doi:10.1128/JB.05017-11

Complete Genome Sequence of the Probiotic Bacterium *Lactobacillus casei* LC2W▽

Chen Chen,[1,2]† Lianzhong Ai,[1]† Fangfang Zhou,[1] Lei Wang,[3] Hao Zhang,[2] Wei Chen,[2]* and Benheng Guo[1]*

State Key Laboratory of Dairy Biotechnology, Technology Center of Bright Dairy & Food Co., Ltd., Shanghai 200436, China[1]; State Key Laboratory of Food Science and Technology, School of Food Science and Technology, Jiangnan University, Wuxi 214122, China[2]; and TEDA School of Biological Sciences and Biotechnology, Nankai University, Tianjin 300457, China[3]

Received 6 April 2011/Accepted 12 April 2011

Lactobacillus casei **LC2W, a patented probiotic strain (Z. Wu, European patent EP 1642963 B1, February 2009), has been isolated from Chinese traditional dairy products and implemented in industrial production as starter culture. Here we present the complete genome sequence of LC2W and the identification of a gene cluster implicated in the biosynthesis of exopolysaccharides.**

Lactobacillus casei LC2W has been isolated from traditional dairy products in Inner Mongolia. Since it possesses the ability to produce exopolysaccharides and probiotic properties (2, 3), this strain has been implemented in industrial production by Bright Dairy & Food Co., Ltd. (8).

Whole-genome sequencing of LC2W was performed with a combined strategy using Roche 454 and Solexa paired-end sequencing technologies. A genomic library containing an 8-kb insert was constructed, and 332,990 paired-end reads and 160,331 single-end reads were generated using the GS FLX system, giving 57.21-fold coverage of the genome. The analysis showed that 91.99% of the reads were assembled into 3 large scaffolds, including 118 contigs. A total of 3,020,574 reads (a 3-kb library) were generated to reach a depth of 98.15-fold coverage and mapped to the scaffolds (7). Most of the gaps within the scaffolds were filled by local assembly of 454 and Solexa reads. Remains and the gaps between scaffolds were filled by sequencing PCR products using an ABI 3730 capillary sequencer. Comparative genomic analysis was performed with the published *L. casei* genomes of strains ATCC 334, BL23, and Zhang as described previously (5, 6).

The complete genome sequence of LC2W is composed of a circular 3,039,042-bp chromosome and a 38,392-bp plasmid named pLC2W, with mean GC contents of 46.36% and 42.64%, respectively. There are 3,121 coding genes, 5 rRNA operons, and 64 tRNAs in the chromosome and 43 coding genes in the plasmid.

The whole-genome sequence of LC2W has proven its safety and stability in commercial use, with no known pathogenic genes identified. Comparative genome analysis showed that the genomes of LC2W and BL23 are of nearly the same size, which is 0.2 Mb larger than those of the other genomes. Almost all the genes present in LC2W and BL23 exhibit nearly identical sequences except for some differential genes. In particular, an insertion island (LCABL_12890 to LCABL_13480) was exclusive for BL23, and the majority of genes were predicted to encode prophage-related proteins, suggesting the region is a prophage remnant for BL23.

LC2W is able to synthesize several fractions of exopolysaccharide (EPS) from skim milk (3), and EPS-related genes were found in its genome. In particular, a 20.4-kb gene cluster (LC2W_2169 to LC2W_2189) which contains 17 EPS-related genes was found in the chromosome, with typical genetic organization and structure for the biosynthesis of EPS. These genes are involved in the regulation, polymerization, and chain length determination and export of the EPS (4). The glycosyltransferase genes in the cluster suggested that the EPS may be composed of rhamnose, glucose, and galactose for the heteropolysaccharide repeating unit with partial acetyl substitution, which is in accordance with the possible composition of the main fraction of the EPS for LC2W, named LCP1 (1), suggesting that the cluster is responsible for biosynthesis for LCP1.

Nucleotide sequence accession numbers. Genome information for the chromosome and plasmid of *Lactobacillus casei* BD-II has been deposited in the GenBank database with accession numbers CP002616 and CP002617.

This work was supported by National Basic Research Program of China (2010CB735705), Shanghai Rising-Star (follow-up-supported) Program (11QH1400700), National Natural Science Foundation of China (30871952 and 20836003), and National High Technology Research and Development Program of China (2010AA1000693002).

* Corresponding author. Mailing address for Benheng Guo: State Key Laboratory of Dairy Biotechnology, Technology Center of Bright Dairy & Food Co., Ltd., 228 Jiangchangsan Road, Shanghai 200436, China. Phone: 86-021-66308993. Fax: 86-021-66308992. E-mail: guobenheng@brightdairy.com. Mailing address for Wei Chen: State Key Laboratory of Food Science and Technology, School of Food Science and Technology, Jiangnan University, 1800 Lihu Road, Wuxi 214122, China. Phone: 86-0510-85912123. Fax: 86-0510-85913583. E-mail: weichen@jiangnan.edu.cn.

† These authors contributed equally to this work.

▽ Published ahead of print on 22 April 2011.

REFERENCES

1. **Ai, L.** 2007. Study on preparation, function and structure of exopolysaccharides from Lactobacillus casei LC2W. Ph.D. thesis. Jiangnan University, Wuxi, China.
2. **Ai, L., et al.** 2008. Isolation and antihypertensive effect of exopolysaccharides from Lactobacillus casei LC2W. Milchwissenschaft **63:**3–6.
3. **Ai, L., et al.** 2008. Preparation, partial characterization and bioactivity of exopolysaccharides from Lactobacillus casei LC2W. Carbohydr. Polym. **74:** 353–357.
4. **Dabour, N., and G. LaPointe.** 2005. Identification and molecular character-

ization of the chromosomal exopolysaccharide biosynthesis gene cluster from Lactococcus lactis subsp. cremoris SMQ-461. Appl. Environ. Microbiol. **71:** 7414–7425.
5. **Feng, L., et al.** 2008. A recalibrated molecular clock and independent origins for the cholera pandemic clones. PLoS One **3:**e4053.
6. **Ferenci, T., et al.** 2009. Genomic sequencing reveals regulatory mutations and recombinational events in the widely used MC4100 lineage of Escherichia coli K-12. J. Bacteriol. **191:**4025–4029.
7. **Li, H., and R. Durbin.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:**1754–1760.
8. **Wu, Z., et al.** February 2009. Lactobacillus casei LC2W strain and its use in antihypertensive aspect. European patent EP 1642963 (B1).