**Genome Sequence of *Lactobacillus rhamnosus* Strain CASL, an Efficient L-Lactic Acid Producer from Cheap Substrate Cassava**

Bo Yu,1,* Fei Su,2 Limin Wang,1 Bo Zhao,1 Jiayang Qin,3 Cuiqing Ma,3 Ping Xu,2,* and Yanhe Ma1

Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, People’s Republic of China; State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, People’s Republic of China; and State Key Laboratory of Microbial Technology, Shandong University, Jinan 250100, People’s Republic of China

Received 29 September 2011/Accepted 30 September 2011

*Lactobacillus rhamnosus* is a type of probiotic bacteria with industrial potential for l-lactic acid production. We announce the draft genome sequence of *L. rhamnosus* CASL (2,855,156 bp with a G+C content of 46.6%), which is an efficient producer of l-lactic acid from cheap, nonfood substrate cassava with a high production titer.

---

**Lactobacillus rhamnosus** is a type of probiotic bacteria and was first isolated in 1983, when it was shown to have remarkable tolerance for harsh acids (1). It was originally considered to be a subspecies of *Lactobacillus casei*, but later genetic research found it to be a species of its own. Lactic acid, the most important hydroxycarboxylic acid, is also a building block for biodegradable plastic, mainly poly l-lactic acid. Therefore, the demand for l-lactic acid is continuously increasing (2, 8). Because of their high levels of l-lactic acid production, some *L. rhamnosus* strains have been suggested to be good producers with great industrial potential (9).

Most studies on l-lactic acid production have focused on the use of pure and easily fermentable sugars (5). Due to the high costs of these pure substrates, the processes are less economical for industrial applications. l-Lactic acid production costs might be significantly reduced if cheap raw materials could be used, such as cassava (4) and Jerusalem artichoke (3). Cassava is one of the most efficient crops in terms of carbohydrate production. It is a tropical perennial plant that grows on poor or depleted soils in which the yields of other crops are very low (10). *L. rhamnosus* CASL is a homofermentative l-lactic acid producer, and the optical purity (enantiomeric excess [ee]) of the l-lactic acid produced is above 98% (10). This strain can also produce a high concentration of l-lactic acid (175 g/liter) from cassava powder in a simultaneous saccharification and fermentation process. This is the highest l-lactic acid concentration reported from a cassava source, and it provides an efficient l-lactic acid production process with this cheap raw bioresource (9). The strain has been deposited in the China General Microbiological Culture Collection Center (CGMCC no. 2183).

Here we announce the draft genome sequence of *L. rhamnosus* obtained by using the Illumina GA system with a paired-end library. The reads were assembled using Velvet (11). The genome sequence of *L. rhamnosus* CASL was annotated with Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (7). In addition, the contigs were searched against the KEGG, UniProt, and Clusters of Orthologous Groups (COG) databases to annotate the gene descriptions. The G+C moles percent measurement was calculated using the draft genome sequence.

The draft genome sequence (2,855,156 bp) contains 168 contigs; the length of the longest contig is 107,604 bp. The GC content of the draft genome sequence is 46.6%, with 3,057 predicted coding sequences. Most genes encoding proteins responsible for glycolysis, amino acid and nucleotide formation, and lipid metabolism were successfully annotated. The obtained genome sequence provides useful hints for strain improvement; for example, the several annotated lactate dehydrogenase genes provide targets for gene knockout to further improve the optical purity of final l-lactic acid production by *L. rhamnosus* CASL.

**Nucleotide sequence accession number.** The whole-genome shotgun sequencing project has been deposited in DDBJ, EMBL, and GenBank under the accession number AFYD00000000. The version described in this paper is the first version, AFYD01000000.

This work was partly supported by grants from the National Basic Research Program of China (2011CBA00806), the Chinese National Programs for High Technology Research and Development (2011AA02A202), the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EW-G-2-3), and the National Natural Science Foundation of China (30900022 and 30821005).

We acknowledge Helmut Bloecker and his colleagues for genome sequencing at the Helmholtz Center for Infection Research in Germany.

**REFERENCES**


