Genome Sequence of *Leuconostoc fallax* KCTC 3537

Seong-Hyeuk Nam,† Sang-Haeng Choi,† Aram Kang,‡ Dong-Wook Kim,† Dae-Soo Kim,† Ryong Nam Kim,† Aeri Kim,‡ and Hong-Seog Park‡,*

Genome Resource Center, Korea Research Institute of Bioscience and Biotechnology (KRIIB), 111 Gwahangno, Yuseong-gu, Daejeon 305-806, Republic of Korea,† and University of Science and Technology (UST), 113 Gwahangno, Yuseong-gu, Daejeon 305-333, Republic of Korea‡

Received 28 October 2010/Accepted 4 November 2010

**Leuconostoc fallax** is known to be present during the manufacturing process of kimchi, the best-known traditional Korean dish. Here, we present the draft genome sequence of the type strain *Leuconostoc fallax* KCTC 3537 (1,638,971 bp, with a G+C content of 37.5%), which consists of 30 large contigs (>100 bp in size).

Kimchi is known to be an important source of vitamins, minerals, and dietary fiber, as well as a good dietary source of lactic acid bacteria (5), and it has been recognized as a health-promoting functional food in recent years (12). Our laboratory received the *Leuconostoc fallax* KCTC 3537 strain, which is known to be present in kimchi (2), from the Korean Collection for Type Cultures (KCTC), and it was grown under standard conditions (lactobacillus MRS Broth [Difco catalog no. 0881], 30°C and 200 rpm). The genomic DNA was extracted from the cultured bacteria using the alkaline lysis method (4). We then sequenced the genome of *Leuconostoc fallax* KCTC 3537; genome sequencing of this organism had not been completed or initiated when our sequencing project was begun, according to the Genomes OnLine Database (GOLD) (10).

Here, we report the genome sequence of *Leuconostoc fallax* KCTC 3537 obtained using a whole-genome shotgun strategy (7), using Roche 454 GS (FLX titanium) pyrosequencing (159,380 reads totaling ~42 Mb; ~26-fold coverage of the genome) at the Genome Resource Center, KRIIB (Korea Research Institute of Bioscience and Biotechnology). Genome sequences from pyrosequencing were processed with Roche’s software according to the manufacturer’s instructions. All of the reads were assembled using Newbler Assembler 2.3 (454 Life Science), which generated 30 large contigs (>100 bp in size; GenBank sequence accession numbers AEIZ01000001 to AEIZ01000030). The annotation was done by merging the results obtained from the RAST (Rapid Annotation using Subsystem Technology) server (1), Glimmer 3.02 modeling software package (6), tRNAscan-SE 1.21 (11), and Rfam 3.2 (9). In addition, the contigs were searched against the KEGG (8), UniProt (3), and Clusters of Orthologous Groups (COG) (13) databases to annotate the gene descriptions. The G+C mole percent measurements were calculated using the genome sequence. The uncompleted draft genome includes 1,638,971 bases and is comprised of 1,895 predicted coding sequences (CDSs), with a G+C content of 37.5%. There are single predicted copies of the 5S, 16S, and 23S rRNA genes and 50 predicted tRNAs. There are 210 subsystems represented in the genome, and we used this information to reconstruct the metabolic network (determined using the RAST server). There are many protein metabolism subsystem features, including protein biosynthesis machinery such as 32 large subunits (LSU) and 21 small subunits (SSU) of the bacterial ribosome and universal GTPases. There are also carbohydrate subsystem features, including genes involved in central carbohydrate, monosaccharide, and fermentation metabolism. There are four predicted d-lactate dehydrogenase enzymes (EC 1.1.1.28) and two alcohol dehydrogenase enzymes (EC 1.1.1.1). The CDSs annotated by COG were classified into 4 COG categories (L, K, R, and S) and 10 COGs (COG0488, COG0536, COG0779, COG1272, COG1399, COG1939, COG2739, COG2740, COG2827, and COG4123). A more detailed analysis of this genome and comparative analysis with other *Leuconostocaceae* genomes will provide further insight into the genomic differences among different species and into *Leuconostocaceae* metabolism.

**Nucleotide sequence accession numbers.** This Whole Genome Shotgun project has been deposited at GenBank under the accession number AEIZ00000000. The version described in this paper is the first version, AEIZ01000000. The 30 large contigs contained in the genome have been deposited under accession numbers AEIZ01000001 through AEIZ01000030.

This work was supported by grant no. 2009-0084206 from the Ministry of Education, Science and Technology.

We thank Kun-Hyang Park and Min-Young Kim for their work in sequencing and assembling the genome, respectively.

**REFERENCES**


