

Draft genome sequence of oligosaccharide producing *Leuconostoc lactis* CCK940 isolated from kimchi in Korea

Sulhee Lee^{ID} and Young-Seo Park*^{ID}

Department of Food Science and Biotechnology, Gachon University, Seongnam 13120, Republic of Korea

올리고당을 생산하는 *Leuconostoc lactis* CCK940 균주의 유전체 염기서열

이설희^{ID} · 박영서*^{ID}

가천대학교 식품생물공학과

(Received September 7, 2018; Revised October 3, 2018; Accepted October 5, 2018)

Leuconostoc lactis CCK940, which was isolated from kimchi obtained from a Korean traditional market, produced an oligosaccharide with a degree of polymerization of more than 4. In this study, the draft genome sequence of *L. lactis* CCK940 was reported by using PacBio 20 kb platform. The genome of this strain was sequenced and the genome assembly revealed 2 contigs. The genome was 1,741,511 base pairs in size with a G + C content of 43.33%, containing 1,698 coding sequences, 12 rRNA genes, and 68 tRNA genes. *L. lactis* CCK940 contained genes encoding glycosyltransferase, sucrose phosphorylase, maltose phosphorylase, and β -galactosidase which could synthesize oligosaccharide.

Keywords: *Leuconostoc lactis*, genome sequence, lactic acid bacteria, oligosaccharide, PacBio

The genus *Leuconostoc* was first described by van Tieghem (1878), and its type species is *Leuconostoc mesenteroides*. The *Leuconostoc* species are Gram-positive, non-motile, and catalase-negative microorganisms with spherical morphology. A recent study reported 14 species of *Leuconostoc*—*L. mesenteroides*, *L. lactis*, *L. citreum*, *L. kimchi*, *L. pseudomesenteroides*, *L. carnosum*, *L. gelidum*, *L. fallax*, *L. gasicomitatum*, *L. garlicum*, *L. inhae*, *L. palmae*, *L. miyukkimchii*, and *L. holzafelii* (Hemme,

2012). *L. argentinum* has been reclassified as a later synonym of *L. lactis* because of their high relatedness (Vancanneyt *et al.*, 2006). The genome size varies between 1.6 Mb and 2.1 Mb, consisting of 30 to 98 contigs (> 100 base pairs [bp] in size), and coding sequences from 1,774 to 2,205 bp (Hemme, 2012). Additionally, the G + C content of the DNA ranges from 36 to 43% (Hemme, 2012).

L. lactis CCK940, which was isolated from kimchi obtained in Gangwon-do, South Korea (37°20'N, 127°56'E), produced oligosaccharide via an acceptor-donor reaction using sucrose as a donor molecule and maltose as an acceptor molecule (Lee and Park, 2017). Despite the health beneficial role of oligosaccharide produced by *L. lactis*, few studies have examined with *L. lactis* at the genome level. Thus, a genome sequence of *L. lactis* CCK940 was determined and its genomic features of the oligosaccharide production in this strain that was related to carbohydrate properties was analyzed.

The strain was cultured in MRS medium at 37°C for 24 h, and genomic DNA was extracted using an AccuPrep[®] Genomic DNA Extraction kit (Bioneer). The genomic DNA of *L. lactis* CCK940 was sequenced by ChunLab, Inc. using the PacBio 20 kb platform. The PacBio reads were assembled using PacBio SMRT analysis pipeline version 2.3.0, and gene expectation was analyzed using CLgenomics 1.55 software (ChunLab, Inc.).

*For correspondence. E-mail: ypark@gachon.ac.kr;
Tel.: +82-31-750-5378; Fax: +82-31-750-5389

The PacBio platform provided $337.1 \times$ coverage of the genome, which was assembled *de novo* into 2 contigs. When chimera test was carried out using Pintail 1.1 program (Ashelford *et al.*, 2005), it was confirmed that there was no sequence contamination with closely related strains. Gene prediction and the basic annotation for the assembled genome were performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). Coding sequences (CDSs) were detected using GenMarkS+ (Besemer *et al.*, 2001). Functional gene annotations of the genome were performed against various databases, including the clusters of orthologous groups (COG) (Tatusov *et al.*, 2000) and NCBI reference sequence (RefSeq) (O'Leary *et al.*, 2016).

Genome sequences of seven strains belonging to *L. lactis* were obtained from the EzGenome database (<https://www.ezbiocloud.net/>) and compared with that of *L. lactis* CCK940 (Table 1). The genome of *L. lactis* CCK940 was 1,741,511 bp in size with a G + C content of 43.33%. The number of protein-coding genes, ribosomal RNA, and transfer RNA genes were 1,698, 12, and 68, respectively. The genome consisted of 353 genes involved in information storage and processing, 266 genes for cellular processes, and 498 genes for cell metabolism, while the functions of 396 genes were not identified. *L. lactis* EFEL005 has 143 ORFs in R (general function prediction only) (Moon *et al.*, 2015), while *L. lactis* CCK940 does not have any ORF in R. Additionally, 21 CRISPR spacers were predicted by CRISPRfiner (Grissa *et al.*, 2007), and the direct repeats were GTACTCGAACTCATTAAAATGACAGACATCTGAAGC (36 bp). *L. lactis* CCK940 contained several genes capable of synthesizing oligosaccharides such as genes encoding β -galactosidase (EC 3.2.1.23), sucrose phosphorylase (EC 2.4.1.7), and maltose phosphorylase (EC 2.4.1.8), which belonged to the GH1, GH13, and GH65 families, respectively. These enzymes, which belong to glucosyltransferase, can produce oligosaccharides

Table 1. General genomic features of *L. lactis* CCK940

Features	Chromosome
Genome size (bp)	1,741,511
Contigs	2
GC content (%)	43.33
rRNA genes	12
tRNA genes	68
Protein coding genes	1,698
Genes assigned to COGs	1,513

of which degree of polymerization is 3 to 9 via an acceptor-donor reaction. Therefore, one of genes encoding these glucosyltransferases would be responsible for the production of oligosaccharide from *L. lactis* CCK940.

L. lactis CCK940 was deposited in Korean Culture Center of Microorganisms (Seoul, Korea) under the deposit number of KFCC11724P.

Nucleotide sequence accession number

This whole-genome shotgun sequencing project has been deposited in GenBank under the accession number NZ_NQLF00000000 for *Leuconostoc lactis* CCK940.

적 요

한국의 전통시장에서 구입한 김치에서 분리된 *Leuconostoc lactis* CCK940은 sucrose와 maltose를 이용하여 중합도가 4 이상인 올리고당을 생산하였다. *L. lactis* CCK940의 유전체는 1,741,511 bp의 2개 contig로 구성된 염색체로 조합되었으며 G + C의 비율은 43.33%로 나타났다. 염색체 DNA에서 1,698개의 코딩 유전자, 12개의 rRNA, 68개의 tRNA 유전자가 확인되었다. *L. lactis* CCK940은 올리고당을 생산할 수 있는 sucrose phosphorylase, maltose phosphorylase, β -galactosidase 등의 glucosyltransferase 생합성 유전자들을 지니고 있었다.

Acknowledgements

This study was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPET) through the High Value-added Food Technology Development Program, funded by the Ministry of Agriculture, Food and Rural Affairs (MAFRA) (Grant number: 316051-03-2-HD020).

References

- Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, and Weightman AJ. 2005. At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Appl. Environ. Microbiol.* **71**, 7724–7736.

- Besemer J, Lomsadze A, and Borodovsky M.** 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* **29**, 2607–2618.
- Grissa I, Vergnaud G, and Pourcel C.** 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **35**, 52–57.
- Hemme D.** 2012. *Leuconostoc* and its use in dairy technology, pp. 73–107. In Hui YH and Evranuz EO. (eds.), Handbook of animal-based fermented food and beverage technology, 2nd ed. Taylor & Francis Group, Boca Raton, Florida, USA.
- Lee S and Park YS.** 2017. Oligosaccharide production by *Leuconostoc lactis* CCK940 which has glucansucrase activity. *Food Eng. Prog.* **21**, 383–390.
- Moon JS, Choi HS, Shin SY, Noh SJ, Jeon CO, and Han NS.** 2015. Genome sequence analysis of potential probiotic strain *Leuconostoc lactis* EFEL005 isolated from kimchi. *J. Microbiol.* **53**, 337–342.
- O'Leary NA, Wright MW, Brister JR, Ciufu S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al.** 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **4**, 733–745.
- Tatusov RL, Galperin MY, Natale DA, and Koonin EV.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36.
- Vancanneyt M, Zamfir M, De Wachter M, Cleenwerck I, Hoste B, Rossi F, Dellaglio F, De Vuyst L, and Swings J.** 2006. Re-classification of *Leuconostoc argentinum* as a later synonym of *Leuconostoc lactis*. *Int. J. Syst. Evol. Microbiol.* **56**, 213–216.
- van Tieghem P.** 1878. Sur la gomme du sucre (*Leuconostoc mesenteroides*). *Ann. Sci. Nat. Bot.* **7**, 180–203.