

Complete genome sequence of *Lactococcus lactis* strain K_LL005, a xylose-utilizing bacterium isolated from grasshopper (*Oxya chinensis sinuosa*)

Hyeri Kim^{1#}, Robin B. Guevarra^{1#}, Jae Hyoung Cho^{1#}, Hyeun Bum Kim^{1*} and Ju-Hoon Lee^{2*}

¹Department of Animal Resources Science, Dankook University, Cheonan 31116, Korea

²Department of Food and Animal Biotechnology, Department of Agricultural Biotechnology, Center for Food and Bioconvergence, Seoul National University, Seoul 08826, Korea



Received: Sep 10, 2020

Revised: Nov 6, 2020

Accepted: Nov 10, 2020

#These authors contributed equally to this work.

*Corresponding author

Hyeun Bum Kim

Department of Animal Resources Science, Dankook University, Cheonan 31116, Korea.

Tel: +82-41-550-3653

E-mail: hbkim@dankook.ac.kr

Ju-Hoon Lee

Department of Food and Animal Biotechnology, Department of Agricultural Biotechnology, Center for Food and Bioconvergence, Seoul National University, Seoul 08826, Korea
Tel: +82-2-880-4854
E-mail: juhlee@snu.ac.kr

Copyright © 2021 Korean Society of Animal Sciences and Technology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Lactococcus lactis is a fermentative lactic acid bacterium that is used extensively in food fermentations. The *L. lactis* strain K_LL005 was isolated from the grasshopper (*Oxya chinensis sinuosa*) gut in Korea. In this study, we reported the complete genome sequence of *Lactococcus lactis* K_LL005. The final complete genome assembly consist of one circular chromosome (2,375,093 bp) with an overall guanine + cytosine (G + C) content of 35.0%. Annotation results revealed 2,281 protein-coding sequences (CDSs), 19 rRNAs, and 68 tRNA genes. *Lactococcus lactis* K_LL005 has a gene encoding xylose metabolism such as *xyIR*, *xyIA*, and *xyIB* (*xyIRAB*).

Keywords: *Lactococcus lactis* strain K_LL005, Grasshopper, Whole genome sequencing, Xylose

In nature, *Lactococcus lactis* was originally isolated from plant or animal surfaces, where it was known to be dormant, and only actively multiplying in the intestinal tract after being consumed by a ruminant [1]. For the past 20 years, *L. lactis* has extremely extended its application from food to being a successful microbial industry. As a result, the species of *L. lactis* has been domesticated for centuries in the fermentation of food, especially cheese and yoghurt [2].

In the present study, the *L. lactis* strain K_LL005 was isolated from the gut of a grasshopper (*Oxya chinensis sinuosa*) collected from local farm in Yangyang, Gangwon-do, Korea. The *L. lactis* K_LL005 was grown in de Man-Rogosa-Sharpe broth at 37 °C for 24 h. Genomic DNA was extracted using the MagAttract HMW DNA Kit (QIAGEN, Hilden, Germany), according to the manufacturer's instructions. The complete genome of the *L. lactis* K_LL005 was sequenced using the PacBio RS II (Pacific Biosciences, Menlo Park, CA, USA) platform at Insilicogen (Yongin, Korea). Library preparation was performed using SMRTbell™ Template Prep Kit 1.0 following the manufacturer's instructions (Pacific Biosciences). PacBio sequencing resulted in 798,317,798

ORCID

Hyeri Kim
<https://orcid.org/0000-0002-6560-2390>
 Robin B. Guevarra
<https://orcid.org/0000-0002-4809-1816>
 Jae Hyoung Cho
<https://orcid.org/0000-0002-1128-3451>
 Hyeun Bum Kim
<https://orcid.org/0000-0003-1366-6090>
 Ju-Hoon Lee
<https://orcid.org/0000-0003-0405-7621>

Competing interests

No potential conflict of interest relevant to this article was reported.

Funding sources

The present study was conducted by the research fund of Dankook University (R201901332).

Acknowledgements

We thank Mo Re Kim (Brandeis University, Waltham, MA, USA) for the English grammar corrections.

Availability of data and material

The complete genome sequences of *Lactococcus lactis* K_LL005 were deposited in GeneBank under the accession numbers CP060580. The BioSample accession number is SAMN15183971, and BioProject accession number is PRJNA638236.

Authors' contributions

Conceptualization: Kim HB, Lee JH.
 Data curation: Kim H, Cho JH.
 Formal analysis: Kim H, Guevarra RB, Cho JH.
 Methodology: Kim H, Guevarra RB, Cho JH.
 Software: Kim H, Guevarra RB, Cho JH.
 Validation: Kim H, Guevarra RB, Cho JH.
 Investigation: Lee JH, Kim HB.
 Writing - original draft: Kim H, Guevarra RB, Cho JH, Kim HB.
 Writing - review & editing: Kim H, Guevarra RB, Cho JH, Kim HB, Lee JH.

Ethics approval and consent to participate

This article does not require IRB/IACUC approval because there are no human and animal participants.

Table 1. Genome features of *Lactococcus lactis* strain K_LL005

Property	Term
Libraries used	PacBio SMRTbell™ library
Sequencing platforms	PacBio RS II sequencer
Assemblers	PacBio SMRT analysis v2.3.0 HGAP.3
Annotation method	Glimmer v3.02, PROKKA v1.14.5 and RAST v2.0
Average genome coverage	267x
Chromosome length (bp)	2,375,093 bp
No. of contigs	1
Guanine + cytosine (G + C) content (%)	35.0
Protein-coding genes (CDSs)	2,281
rRNA genes	19
tRNA genes	68
Plasmids	0
Genbank Accession No.	CP060580

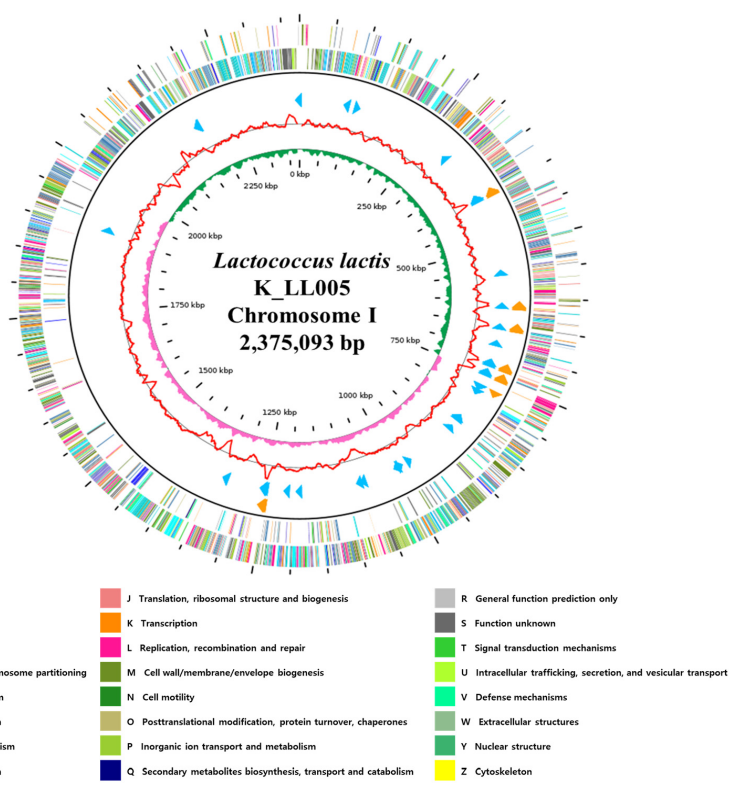


Fig. 1. Genome map of *Lactococcus lactis* strain K_LL005. The outer circle denotes the locations of all annotated ORFs, and the inner circle with the red denotes GC content. Pink and green peaks denote GC skew. The orange arrows denote rRNAs, and the sky blue arrows denote the tRNA operons. All annotated ORFs are colored differently based on the COG assignments. G, guanine; C, cytosine; ORFs, open reading frames; COG, clusters of orthologous groups.

base pairs and 109,112 long reads after filtering. The sequence reads were *de novo* assembled using the hierarchical genome assembly process (HGAP v2.3.0) workflow, and further polished with Quiver. Because bacterial genomes and plasmids are typically circular, we checked the forms for each of contigs using MUMmer 3.5 [3] and trimmed one of the self-similar

ends for manual genome closure. Putative protein coding sequences were predicted using Glimmer v3.02 [4], rapid prokaryotic genome annotation (PROKKA) v1.14.5 [5] and Rapid Annotation using Subsystem Technology (RAST) v2.0 [6]. The results from the three prediction methods were confirmed by manual curation [7]. The genes of rRNA and tRNA were predicted using RNAmmer v1.2 and tRNAscan-SE v1.4, respectively [8,9]. GO annotation was assigned to each of ORFs using Blast2GO software [10] to analyze the best hits of the BLAST results. The complete genome of the *L. lactis* strain K_LL005 contains one circular chromosome (2,375,093 bp) with a guanine + cytosine (GC) content of 35.0%, 2281 predicted open reading frames (ORFs), 19 rRNA genes, and 68 tRNA genes. The genome feature and map of *L. lactis* strain K_LL005 are illustrated in Table 1 and Fig 1.

The genome of *L. lactis* K_LL005 possessed *xylA*, *xylB*, and *xylR* (*xylRAB*) encoding enzymes and protein for xylose utilization. Xylose is a component of hemicellulose, an important structural polysaccharide in plants, and xylose utilization metabolism required the products of the *xylRAB* genes encoding putative positive transcriptional regulator, xylose isomerase, and xylulose kinase. Furthermore, *L. lactis* K_LL005 has the genes, *MnnB*, *RbsK*, and *FrcK*, which are associated with mannose utilization, D-ribose utilization and fructose utilization.

REFERENCES

1. Guevarra RB, Lee JH, Lee SH, Seok MJ, Kim DW, Kang BN, et al. Piglet gut microbial shifts early in life: causes and effects. *J Anim Sci Biotechnol*. 2019;10:1. <https://doi.org/10.1186/s40104-018-0308-3>
2. Song AA, In LLA, Lim SHE, Rahim RA. A review on *Lactococcus lactis*: from food to factory. *Microb Cell Factories*. 2017;16:55. <https://doi.org/10.1186/s12934-017-0669-x>
3. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. *Genome Biol*. 2004;5:R12. <https://doi.org/10.1186/gb-2004-5-2-r12>
4. Delcher AL, Bratke KA, Powers EC, Salzberg SL. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*. 2007;23:673-9. <https://doi.org/10.1093/bioinformatics/btm009>
5. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068-9. <https://doi.org/10.1093/bioinformatics/btu153>
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. *BMC Genom*. 2008;9:75. <https://doi.org/10.1186/1471-2164-9-75>
7. Lee JE, Heo S, Kim GB. Complete genome sequence of *Streptococcus hyointestinalis* B19, a strain producing bacteriocin, isolated from chicken feces. *J Anim Sci Technol*. 2020;62:420-2. <https://doi.org/10.5187/jast.2020.62.3.420>
8. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;35:3100-8. <https://doi.org/10.1093/nar/gkm160>
9. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25:955-64. <https://doi.org/10.1093/nar/25.5.955>
10. Conesa A, Götz S, García-Gomóez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21:3674-6. <https://doi.org/10.1093/bioinformatics/bti610>