JAST (Journal of Animal Science and Technology) TITLE PAGE
Upload this completed form to website with submission

ARTICLE INFORMATION	**Fill in information in each box below**
Article Type	Short Communication
Article Title (within 20 words without abbreviations)	Complete genome sequence of functional probiotic candidate *Lactobacillus amylovorus* CACC736
Running Title (within 10 words)	Genome sequence analysis of *Lactobacillus amylovorus* CACC736
Author	Soyeon Park[1], Jung-Ae Kim[1], Hyun-Joon Jang[1], Dae-Hyuk Kim[1,2], and Yangseon Kim[1]
Affiliation	1Department of Research and Development, Center for Industrialization of Agricultural and Livestock Microorganisms, Jeongeup, 56212, Korea 2Department of Molecular Biology, Department of Bioactive Material Science, Institute for Molecular Biology and Genetics, Jeonbuk National University, Jeonju, 54896, Korea
ORCID (for more information, please visit https://orcid.org)	Soyeon Park (http://orcid.org/0000-0003-3788-5415) Jung-Ae Kim (http://orcid.org/0000-0002-0694-477x) Hyun-Jun Jang (https://orcid.org/0000-0003-2906-7543) Dae-Hyuk Kim (http://orcid.org/0000-0002-9948-5313) Yangseon Kim (http://orcid.org/0000-0002-8285-3407)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	This research was supported by INNOPOLIS FOUNDATION through Science and Technology Project Opens the Future of the Region, which is funded by Ministry of Science and ICT (2022-DD-UP-0333), Korea, and partially supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through Useful Agricultural Life Resources Industry Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (321094-2).
Acknowledgements	Not applicable.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form.	Conceptualization: Kim Y. Data curation: Park S, Jang HJ. Formal analysis: Park S. Methodology: Park S, Kim JA. Software: Park S. Validation: Park S, Kim DH, Kim JA. Investigation: Kim Y. Writing - original draft: Park S. Writing - review & editing: Kim Y.
Ethics approval and consent to participate	This article does not require IRB/IACUC approval because there are no human and animal participants.

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	**Fill in information in each box below**
First name, middle initial, last name	Yangseon Kim

Email address – this is where your proofs will be sent	yangseon@cialm.or.kr
Secondary Email address	nimitzdr93@gmail.com
Address	241, Cheomdangwahagro, Jeongeup-si, Jeonbuk, Korea
Cell phone number	82-10-5680-4340
Office phone number	82-63-536-6712
Fax number	82-63-536-6003

Abstract

Lactobacillus amylovorus CACC736 was originated from swine feces in Korea. The complete genome sequences of the strain contained one circular chromosome (2,057,809 bp) with 38.2% GC content and two circular plasmids, namely, pCACC736-1 and pCACC736-2. The predicted protein-coding genes, which are encoding the CRISPR-associated proteins, biosynthesis of bacteriocin (helveticin J), and the related proteins of the bile, acid tolerance. Notably, the genes related to vitamin B-group biosynthesis (riboflavin and cobalamin) were also found in *L. amylovorus* CACC736. Collectively, the complete genome sequence of the *L. amylovorus* CACC736 will aid in the development of functional probiotics in the animal industry.

Keywords: ***Lactobacillus amylovorus*, Swine, Probiotics, Whole-genome sequencing**

Announcement

Lactobacillus spp. are non-pathogenic microorganisms that provide beneficial effects to the host [1, 2, 3]. *Lactobacillus amylovorus* has been studied as a paraprobiotic (non-viable cells or cell fractions) with the ability to change body adiposity [1]. Additionally, it has been reported that *L. amylovorus* has probiotic properties such as antiviral and antimicrobial activities through the regulation of the gut microflora [2, 3]. In this study, the genomes of *L. amylovorus* CACC736 are functionally annotated.

L. amylovorus strain CACC736 (KACC22146) was isolated from swine feces in Korea. This strain was inoculated in de Man, Rogosa, and Sharpe (MRS) medium (Difco, Franklin Lakes, NJ, USA) and cultivated at 37 ℃ for 24 h. Genomic DNA (gDNA) of the strain was extracted using the DNeasy UltraClean microbial kit (Qiagen, Hilden, Germany). The complete genome sequence of *L. amylovorus* strain CACC736 was obtained with single-molecule real-time sequencing technology (SMRT) on the platform of PacBio RSII (Pacific Biosciences, Menlo Park, CA, USA) at ChunLab, Inc. (Seoul, Korea). These gene sequences were performed by *de novo* genome assembly using the PacBio SMRT Analysis (version 2.3.0, Pacific Biosciences) [4]. All genes were classified by different functional groups using EggNOG 4.5 (http://eggnog5.embl.de). Additionally, functional annotation of the CDSs was performed by the UBLAST program including the databases of the Swiss-Prot and Kyoto Encyclopedia of Genes and Genomes (KEGG) [5]. Predictions for clustered regularly interspaced short palindromic repeats (CRISPR) were used by CRISPR finder (https://crispr.i2bc.paris-saclay.fr) [6].

The *L. amylovorus* CACC736 composed of one circular chromosome (2,057,809 bp, 38.2% GC content) along with two plasmids designated as pCACC736-1 (76,480 bp, 36.0% GC content) and pCACC736-2 (20,439 bp, 35.0% GC content) (Table 1 and Fig. 1A). Moreover, the complete genome comprised 2,080 protein-coding sequences (CDSs) and 80 non-coding genes (15 rRNA and 65 tRNA genes). A total of 1,848 proteins (88.8%) were classified on a functional categorization by the database of Clusters of Orthologous Groups (COGs) categories (Fig. 1B). The most abundant COGs categories, excluding an ‘unknown function [S]’, were ‘replication, recombination and repair [L]’ (295 genes; 16.0%), ‘carbohydrate transport and metabolism [G]’ (146 genes; 7.9%), ‘translation, ribosomal structure and biogenesis [J]’ (137 genes; 7.4%), and ‘amino acid transport and metabolism [E]’ (126 genes; 6.8%). The genome of the *L. amylovorus* CACC736 encoded five CRISPR genes/proteins (Cas1, Cas2, Cas3, Cas4, and Cas6) for antiviral-relative mechanisms [7], one bacteriocin class III (helveticin J) for an inhibitory effect against common pathogenic organisms [8], and two potential genes of antimicrobial activity (*lysM* and *qac*). In addition, the *L. amylovorus* CACC736 was confirmed to have genes associated with common probiotic properties, such as lactate synthesis (*ldh*, L-lactate dehydrogenase), bile salt hydrolases (BSH; *cbh*) and acid tolerance (*atpD*, *atpH*, and *grpE*) (Table 2). Notably, we revealed the presence of genes involved in vitamin B_2 and B_{12} biosynthesis, including riboflavin (*ribB*, *ribD*, *ribE*, and *ribT*) and cobalamin (*cobC*) (Table 2) [9, 10]. Taken together, our findings on the complete genome of *L. amylovorus* CACC736 will provide a scientific improvement for the development of functional probiotics.

Nucleotide sequence accession numbers

The complete genome sequences of *L. amylovorus* strain CACC736 were deposited at the NCBI GenBank under the accession numbers CP104879 (chromosome) and CP104880-CP104881 (plasmids), respectively.

Acknowledgements

This research was supported by INNOPOLIS FOUNDATION through Science and Technology Project Opens the Future of the Region, which is funded by Ministry of Science and ICT (2022-DD-UP-0333), Korea and partially supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through Useful Agricultural Life Resources Industry Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (321094-2).

References

1. Sugawara T, Sawada D, Yanagihara S, Aoki Y, Takehara I, Sugahara H, et al. Daily intake of paraprobiotic Lactobacillus amylovorus CP1563 improves pre-obese conditions and affects the gut microbial community in healthy pre-obese subjects: a double-blind, randomized, placebo-controlled study. Microorganisms. 2020;8(2):304. https://doi.org/10.3390/microorganisms8020304

2. Martinez RC, Aynaou A-E, Albrecht S, Schols HA, De Martinis EC, Zoetendal EG, et al. In vitro evaluation of gastrointestinal survival of Lactobacillus amylovorus DSM 16698 alone and combined with galactooligosaccharides, milk and/or Bifidobacterium animalis subsp. lactis Bb-12. Int J Food Microbiol. 2011;149(2):152-8. https://doi.org/10.1016/j.ijfoodmicro.2011.06.010

3. Foulquié Moreno MR, Baert B, Denayer S, Cornelis P, De Vuyst L. Characterization of the amylovorin locus of Lactobacillus amylovorus DCE 471, producer of a bacteriocin active against Pseudomonas aeruginosa, in combination with colistin and pyocins. FEMS Microbiol Lett. 2008;286(2):199-206. https://do.org/10.1111/j.1574-6968.2008.01275.x .

4. Kim J-A, Jung MY, Kim D-H, Kim Y. Genome analysis of Bacteroides sp. CACC 737 isolated from feline for its potential application. J Anim Sci Technol. 2020;62(6):952. https://doi.org/10.5187/jast.2020.62.6.952

5. Kang S, Li R, Jin H, You HJ, Ji GE. Effects of selenium-and zinc-enriched Lactobacillus plantarum SeZi on antioxidant capacities and gut microbiome in an ICR mouse model. Antioxidants. 2020;9(10):1028. https://doi.org/10.3390/antiox9101028

6. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007;35(suppl_2):W52-W7. https://doi.org/10.1093/nar/gkm360

7. Rath D, Amlinger L, Rath A, Lundgren M. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie. 2015;117:119-28. https://doi.org/10.1016/j.biochi.2015.03.025

8. Ghodhbane H, Elaidi S, Sabatier J-M, Achour S, Benhmida J, Regaya I. Bacteriocins active against multi-resistant gram negative bacteria implicated in nosocomial infections. Infect Disord Drug Targets. 2015;15(1):2-12. https://doi.org/10.2174/1871526514666140522113337

9. Thakur K, Tomar SK, De S. Lactic acid bacteria as a cell factory for riboflavin production. Microb Biotechnol. 2016;9(4):441-51. https://doi.org/10.1111/1751-7915.12335

10. Capozzi V, Russo P, Dueñas MT, López P, Spano G. Lactic acid bacteria producing B-group vitamins: a great potential for functional cereals products. Appl Microbiol Biotechnol. 2012;96(6):1383-94. https://doi.org/10.1007/s00253-012-4440-2

Tables and Figures

Table 1. General features of *L. amylovorus* CACC736 genome

Properties	**Chromosome**	**Plasmids**	
	CACC736	**pCACC736-1**	**pCACC736-2**
BioProject	PRJNA881772	–	–
BioSample	SAMN30915630	–	–
Accession No.	CP104879	CP104880	CP104881
Genome size (bp)	2,057,809	76,480	20,439
GC content (%)	38.2	36.0	35.0
No. of CDSs	1,989	71	20
No. of CRISPR regions	5	–	–
rRNA genes	15	–	–
tRNA genes	65	–	–

L, Lactobacillus; CDSs, coding sequences; GC, guanine-cytosine; CRISPR, clustered regularly interspaced short palindromic repeats

Table 2. Predicted CDSs involved in probiotic potency in *L. amylovorus* CACC736

Predicted function	***L. amylovorus* CACC736**			
	Predicted genes	**Start position**	**End position**	**Length(bp)**
CRISPR/cas				
Endonuclease	Cas1	c1,507,114	c1,508,103	990
Endonuclease	Cas2	c1,506,827	c1,507,108	282
Endonuclease/helicase	Cas3	c1,508,620	c1,511,049	2,430
Exonuclease	Cas4	c1,508,113	c1,508,604	492
Endoribonuclease	Cas6	c1,515,924	c1,516,679	756
Antimicrobial activity-related				
Lysin Motif domain	*lysM*	890,591	891,055	465
Quaternary ammonium compound-resistance	*qacC*	1,012,642	1,012,962	321
Bacteriocin (Class III)	helveticin J	c1,995,360	c1,995,992	633
Lactate synthesis	*ldh*	1,795,954	1,796,925	972
Bile salt hydrolases (BSH)	*cbh*	1,052,335	1,053,357	1,023
Acid tolerance-related	*atpD*	690,181	691,692	1,512
	atpH	692,699	694,138	1,440
	clpB	98,750	100,879	2,130
	grpE	c1,235,180	c1,235,764	585
Protection or repair-related	*dnaJ*	c1,232,063	c1,233,217	1,155
Vitamin B-groups synthesis				
Vitamin B2	*ribB*	c1,025,985	c1,027,160	1,176
	ribD	c1,027,752	c1,028,810	1,059
	ribE	c1,027,163	c1,027,759	597
	ribT	887,197	887,550	354
Vitamin B12	*cobC*	c301,776	c302,426	651

CDSs, coding sequences; CRISPR, clustered regularly interspaced short palindromic repeats; c, complement

Figure Legends

Figure 1. Genome futures of *Lactobacillus amylovorus* CACC736. (A) Circular genome mapping of *Lactobacillus amylovorus* CACC736. Circles from the outside to the center denote: (a) forward and (b) reverse strands (colored according to COGs function categories), (c) rRNA and tRNA, (d) GC skew, (e) GC content. (B) Functional classification of COGs. COGs, cluster of orthologous groups of proteins database.