Complete genome sequencing and comparative genomic analysis of *Lactobacillus acidophilus* C5 as a potential canine probiotics

Seungwoo Son[1#], Raham Lee[2#], Seung-Moon Park[3], Sung Ho Lee[4], Hak-Kyo Lee[1,2], Yangseon Kim[5]* and Donghyun Shin[1]*

[1]Department of Agricultural Convergence Technology, Jeonbuk National University, Jeonju 54896, Korea
[2]Department of Animal Biotechnology, College of Agricultural and Life Sciences, Jeonbuk National University, Jeonju 54896, Korea
[3]Department of Bioenvironmental Chemistry, College of Agriculture and Life Sciences, Jeonbuk National University, Jeonju 54896, Korea
[4]Woogene B&G, Hwaseong 18630, Korea
[5]Center for Industrialization of Agriculture and Livestock Microorganism, Jeongeup 56212, Korea

*Corresponding author
Yangseon Kim
Center for Industrialization of Agriculture and Livestock Microorganism, Jeongeup 56212, Korea.
Tel: +82-63-536-6001
E-mail: yangseon@cialm.or.kr

Donghyun Shin
Department of Agricultural Convergence Technology, Jeonbuk National University, Jeonju 54896, Korea.
Tel: +82-63-270-5931
E-mail: sdh1214@gmail.com

Abstract

Lactobacillus acidophilus is a gram-positive, microaerophilic, and acidophilic bacterial species. *L. acidophilus* strains in the gastrointestinal tracts of humans and other animals have been profiled, but strains found in the canine gut have not been studied yet. Our study helps in understanding the genetic features of the *L. acidophilus* C5 strain found in the canine gut, determining its adaptive features evolved to survive in the canine gut environment, and in elucidating its probiotic functions. To examine the canine *L. acidophilus* C5 genome, we isolated the C5 strain from a Korean dog and sequenced it using PacBio SMRT sequencing technology. A comparative genomic approach was used to assess genetic relationships between C5 and six other strains and study the distinguishing features related to different hosts. We found that most genes in the C5 strain were related to carbohydrate transport and metabolism. The pan-genome of seven *L. acidophilus* strains contained 2,254 gene families, and the core genome contained 1,726 gene families. The phylogenetic tree of the core genes in the canine *L. acidophilus* C5 strain was very close to that of two strains (DSM20079 and NCFM) from humans. We identified 30 evolutionarily accelerated genes in the *L. acidophilus* C5 strain in the ratio of non-synonymous to synonymous substitutions (dN/dS) analysis. Five of these thirty genes were associated with carbohydrate transport and metabolism. This study provides insights into genetic features and adaptations of the *L. acidophilus* C5 strain to survive the canine intestinal environment. It also suggests that the evolution of the *L. acidophilus* genome is closely related to the host's evolutionary adaptation process.

Keywords: *Lactobacillus acidophilus*, Canine, Comparative genomics, Pan-genome analysis, Host adaptation

ORCID
Seungwoo Son
https://orcid.org/0000-0002-2176-8276
Raham Lee
https://orcid.org/0000-0002-4928-9517
Seung-Moon Park
https://orcid.org/0000-0002-7983-1143
Sung Ho Lee
https://orcid.org/0000-0001-9398-5846
Hak-Kyo Lee
https://orcid.org/0000-0001-5387-4885
Yangseon Kim
https://orcid.org/0000-0002-8285-3407
Donghyun Shin
https://orcid.org/0000-0002-0819-0553

INTRODUCTION

Lactobacillus acidophilus is a gram-positive, homofermentative, and microaerophilic bacteria. It ferments sugars into lactic acid and grows readily in acidic pH (below 5.0) and is common in the gastrointestinal (GI) tract of humans and other animals [1]. *L. acidophilus* is used in the production of fermented food and dairy products and is a symbiont in humans and animals. *L. acidophilus* strains are commercially used in many dairy products for the production of yogurt, health foods, and several medicines [2]. Some strains of *L. acidophilus* have probiotic characteristics. Several studies have shown that *L. acidophilus* exhibits various probiotic effects in humans and animals and helps in lowering cholesterol levels, preventing and treating diarrhea, modulating the immune system, and suppressing cancer [3,4].

Recently, there has been an increasing interest in developing methods to modulate the animal intestinal microbiota to improve health. The GI tract of healthy dogs contains *Lactobacillus* species, such as *L. acidophilus* [5]. Even after considering inter-individual variations, *L. acidophilus* is established in the gut of dogs soon after birth, similar to humans and other mammals. As it grows, it reaches compositional stability, with its principal activity being inhibition of undesirable microorganism proliferation [6]. Additionally, commensal gut bacteria positively interact with the host immune system, producing a wide range of metabolites crucial for host physiology. The gut bacteria depend on their host for nutrients and to maintain a stable ecosystem [7]. Therefore, the phylogenetic differentiation of *L. acidophilus* strains reflects their co-evolution with their vertebrate hosts. Previous studies have characterized the phylogenetic and genetic features of *L. acidophilus* strains isolated from diverse hosts using their genomic data [8, 9]. To date, dozens of strains isolated from several hosts have been studied and are published in the National Center for Biotechnology Information Center (NCBI). However, the phylogenetic and genetic features of *L. acidophilus* strains isolated from canine intestines remain unknown.

Recently, with the development of long-read and high-throughput DNA sequencing technologies, whole-genome studies have become increasingly feasible and affordable, making the genomic data of diverse organisms publicly available. Based on this data, comparative genomic analysis of strains within the same species has provided insights into modified, acquired, or lost genetic features closely related to the evolution and adaptation to specific environments [10]. In this study, we sequenced canine-derived *L. acidophilus* strain (C5) and performed comparative genomic analysis (pan-genome and the ratio of non-synonymous to synonymous substitutions [dN/dS] analysis) to profile its genetic characteristics.

MATERIALS AND METHODS

L. acidophilus C5 isolation

L. acidophilus C5 was isolated from dog feces (Shih Tzu, male) in Korea by Woogene B&G and was kindly provided for research purposes. *L. acidophilus* C5 was cultivated on modified de Man, Rogosa and Sharpe (MRS) (mMRS, with 0.05% cysteine-HCl) in an anaerobic atmosphere (5% hydrogen, 5% carbon dioxide, and 90% nitrogen) for 48 h at 37℃.

Genome sequencing, annotation, and average nucleotide identity (ANI) calculation

DNA was extracted using DNeasy UltraClean Microbial Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Whole-genome shotgun sequencing of the *L. acidophilus* C5 strain was performed using the PacBio SMRT and Illumina HiSeq sequencing technology. For genome assembly, we applied the recently described Unicycler (v0.4.6) [11] and the final assembly was

2.005 Mb in one contig (Fig. 1). The PreAssembly step mapped single-pass reads to seed reads, the longest portion of the read length distribution. Subsequently, a consensus sequence of the mapped reads was generated, resulting in long and highly accurate fragments of the target genome. The next step was to correct and filter the reads. Reads that were fully contained in other reads did not provide extra information for constructing the genome, so they were filtered. Reads with an unsuitable extent of overlap were also filtered. Next, we constructed contigs of the *L. acidophilus* C5 strain. After de-novo assembly, we mapped the Illumina HiSeq reads to the first assembled genome sequence. We observed slight difference in the mapping result and the assembly result. We used this information to generate a consensus sequence of higher quality through a self-mapping step. Previously published genome sequences of six strains of *L. acidophilus* isolated from diverse hosts (yogurt, DuPont nutrition, and humans) were acquired from NCBI and compared with the host-derived strain. All genome sequences of *L. acidophilus* strains were annotated using Prokka (v1.12b) [12] and EggNOG-mapper (v5.0) [13]. The protein-coding sequences were categorized based on the Clusters of Orthologous Groups (COG) database from the Prokka results (Fig. 2). For comparative genomic analysis in this study, we selected strains with complete genome information in NCBI database. To evaluate the genetic relationship between *L. acidophilus* C5 and the other strains, the ANI was calculated using the JSpecies web server [14]. All information on *L. acidophilus* C5 and the other six strains used in this study is presented in Table 1 and Supplementary Table S1.
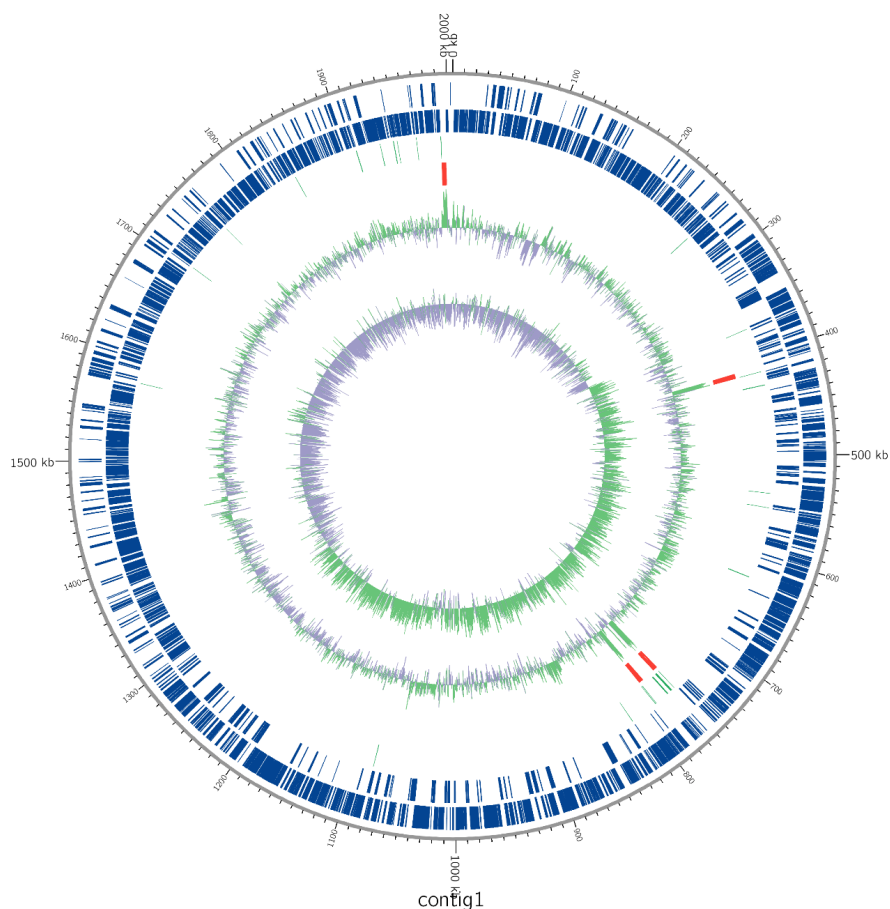

Fig. 1. Genome map of *Lactobacillus acidophilus* C5. Circular map was drawn using the genome annotation result. Marked characteristics are shown from outside to the center; CDS on forward strand, CDS on reverse strand, tRNA, rRNA, GC content, and GC skew. CDS, coding region; G, guanine; C, cytosine.
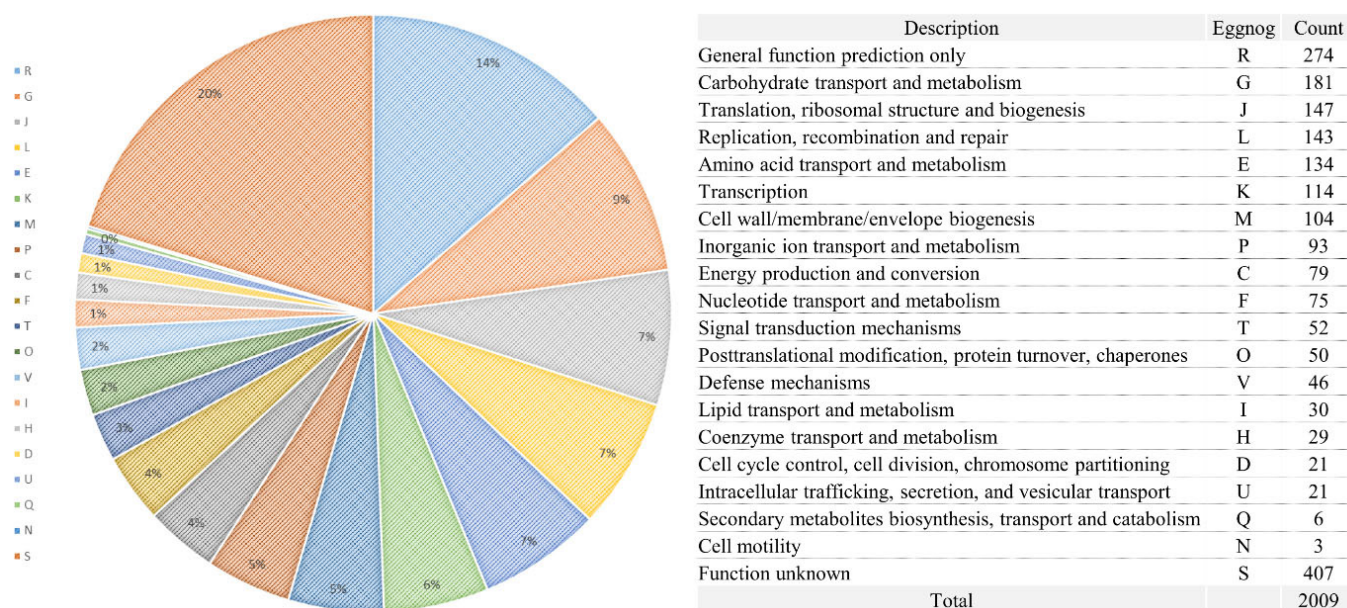
Description	Eggnog	Count
General function prediction only	R	274
Carbohydrate transport and metabolism	G	181
Translation, ribosomal structure and biogenesis	J	147
Replication, recombination and repair	L	143
Amino acid transport and metabolism	E	134
Transcription	K	114
Cell wall/membrane/envelope biogenesis	M	104
Inorganic ion transport and metabolism	P	93
Energy production and conversion	C	79
Nucleotide transport and metabolism	F	75
Signal transduction mechanisms	T	52
Posttranslational modification, protein turnover, chaperones	O	50
Defense mechanisms	V	46
Lipid transport and metabolism	I	30
Coenzyme transport and metabolism	H	29
Cell cycle control, cell division, chromosome partitioning	D	21
Intracellular trafficking, secretion, and vesicular transport	U	21
Secondary metabolites biosynthesis, transport and catabolism	Q	6
Cell motility	N	3
Function unknown	S	407
Total		2009

Fig. 2. Functional categorization using Eggnog annotation (COG database) of all predicted CDSs of *Lactobacillus acidophilus* **C5**. COG, clusters of orthologous groups; CDS, coding region.

Table 1. Comparison of the chromosomal properties of the seven *Lactobacillus acidophilus* strains

Strain	C5	NCFM	La14	FSI4	ATCC53544	LA1	DSM20079
Genome size (bp)	2,005,383	1,993,560	1,991,579	1,991,969	1,991,906	1,991,195	2,009,973
Gene	2,082	1,963	1,978	1,977	1,977	2,002	2,020
Coding genes	2,009	1,832	1,862	1,868	1,852	1,886	1,840
Non coding genes	73	131	116	109	125	116	180
ANI (average nucleotide identity)	100.00%	99.28%	99.28%	99.29%	99.28%	99.28%	99.24%

Pan-genome analysis

The genome sequences of seven *L. acidophilus* strains, including the C5 strain, were first annotated using Prokka (v1.12b) [12] to obtain GFF files, which were used to perform pan-genome analysis. The core- and pan-genomes were calculated using Roary, a rapid standalone pan-genomic pipeline [15] and a pair of genes was defined as belonging to the same gene family when the identity value of their amino acid sequences was > 95% (Fig. 3). The COG annotation (from EggNOG-mapper result) of core genes and the C5 strain is shown in Fig. 4. A phylogenetic tree was constructed based on the core genes (Fig. 5).

Comparing C5 strain to the other strains using dN/dS analysis

For comparative genome analysis using the dN/dS method, we used OrthoFinder (v1.1.10) [16] and PRANK [17] to determine ortholog genes for the seven genomes and multiple sequence alignment of each orthologous gene, respectively. These sequences were converted into the corresponding cDNA sequences using PAL2NAL [18] and poorly aligned transcripts were eliminated using Gblocks [19]. After all the filtering steps, a total of 1,843 orthologous groups remained. Phylogenetic analysis by maximum (PAML4) analysis using the maximum likelihood method [20] was used to estimate the dS and dN. Phylogenetically featured genes were investigated by the branch-site models.
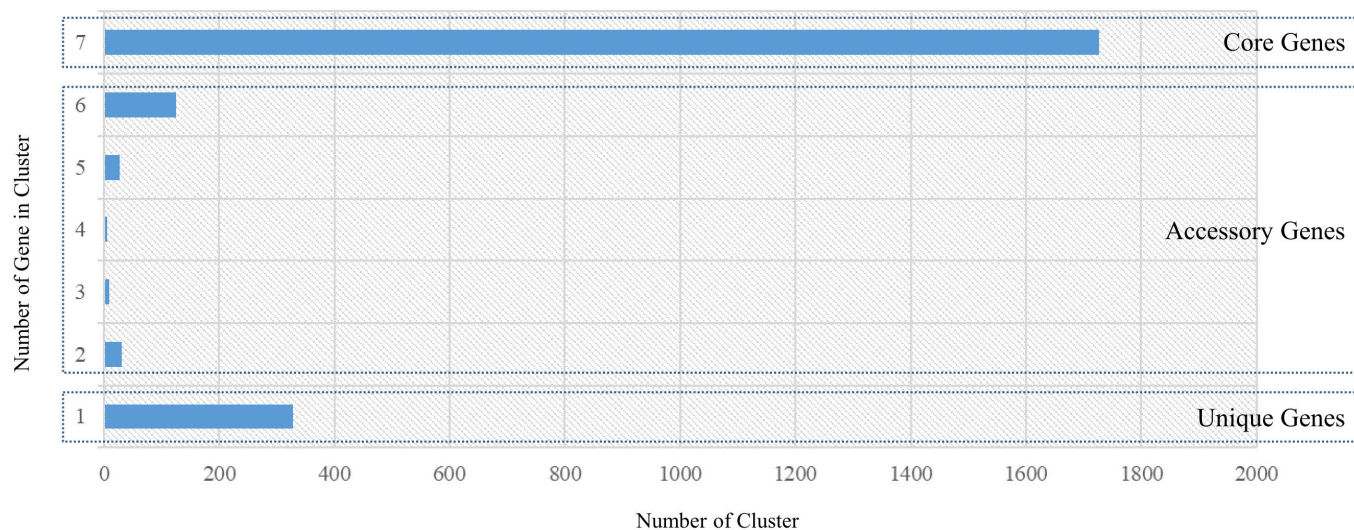
Fig. 3. Pan-genome analysis using genome sequences of seven *Lactobacillus acidophilus* strains. The plot represents the core genes, accessory genes, and unique genes of the seven analyzed genomes (core genes = all seven species, accessory genes = 2–6, and unique genes = 1).
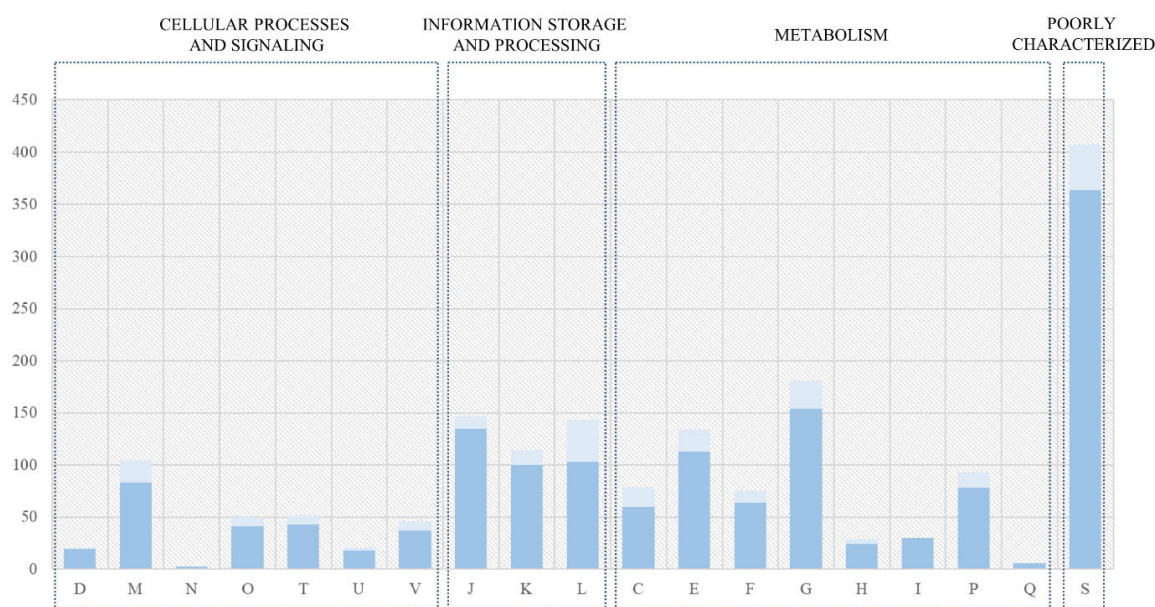

Fig. 4. Core and C5 genes of *Lactobacillus acidophilus* using the Eggnog annotation result (COG database) Blue color means core genes (genes in all strains) and light blue means genes only in C5. Also, the X-axis is information that classifies gene functions by the COG database. The meaning of each abbreviation is as follows. D, cell cycle control, cell division, chromosome partitioning; M, cell wall/membrane/envelope biogenesis; N, cell motility; O, post-translational modification, protein turnover, and chaperones; T, signal transduction mechanisms; U, intracellular trafficking, secretion, and vesicular transport; V, defense mechanisms; J, translation, ribosomal structure, and biogenesis; K, transcription; L, replication, recombination, and repair; C, energy production and conversion; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; P, inorganic ion transport and metabolism; Q, secondary metabolites biosynthesis, transport, and catabolism; S, function unknown; COG, clusters of orthologous groups.

RESULTS

General genomic characteristics of the *Lactobacillus acidophilus* C5 strain

The total genome size of the *L. acidophilus* C5 strain in this study was 2.005 Mb, and the guanine +
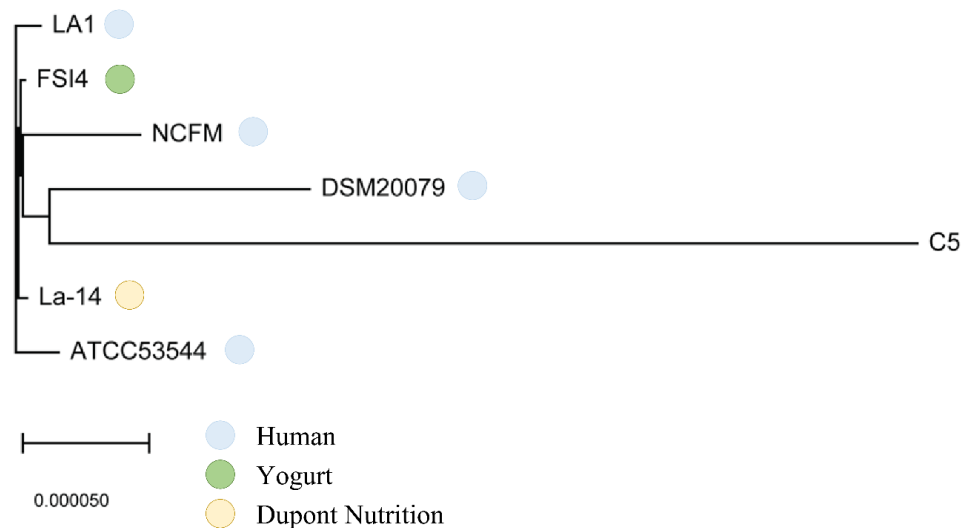
Fig. 5. Phylogenetic analyses (C5 and six other complete genome sequences of *Lactobacillus acidophilus*) using the core genes from pan-genome analysis.

cytosine (G + C) content was 34.5% (after fitting). Additionally, genome annotation using Prokka (v1.12b) [12] and EggNOG-mapper [13] showed that the sequenced genome consisted of 2,009 coding genes and 73 non-coding genes (61 tRNA and 12 rRNA genes) (Fig. 1 and Table 1). In the functional analysis using the COG database in EggNog-mapper, the largest protein-coding categories (except, "General function prediction only" and "Function unknown") in *L. acidophilus* C5 strain were "Carbohydrate transport and metabolism (G)" (9.01%), "Translation, ribosomal structure and biogenesis (J)" (7.32%) and "Replication, recombination, and repair (L)" (7.12%) (Fig. 2).

Pan-genome analysis of *Lactobacillus acidophilus* strains

The core- and pan-genomes of the seven *L. acidophilus* strains, including the C5 strain, were analyzed using a comparative genomics method. The pan-genome of the seven *L. acidophilus* strains contained 2,254 gene families, and the core genome contained 1,726 gene families, indicating that together, the seven genomes were sufficient to represent the core genome of *L. acidophilus*. Moreover, the seven *L. acidophilus* genomes contained 200 accessory gene families (six isolates: 126 gene families, five isolates: 27 gene families, four isolates: 5 gene families, three isolates: 10 gene families, and two isolates: 32 gene families), and 328 strain-specific genes (Fig. 3). To determine the functions of the 1,726 core genes, we extracted the sequences of the core genes to map to the COG database. The results showed that except for the "General function prediction only" and "Function unknown" categories, the largest proportion of core genes belonged to "Carbohydrate transport and metabolism (G, 154 genes)" followed by "Translation, ribosomal structure and biogenesis (J, 135 genes)", and "Amino acid transport and metabolism (E, 113 genes)." (Fig. 4). Using a phylogenetic tree based on core genes, we found that the *L. acidophilus* C5 strain was clearly distinguished from the other six strains (Fig. 5). The strains closest to the *L. acidophilus* C5 strain were the human-derived strains (DSM20079 and NCFM). The C5 strain had the highest number of unique genes (245 genes) among the seven *L. acidophilus* strains. Among the unique genes of the C5 strain, the largest proportion belonged to "Replication, recombination, and repair (L) (28 genes)" and "Carbohydrate transport and metabolism (G) (24 genes)" apart from the "Function unknown" category (Fig. 4).

Comparison of *Lactobacillus acidophilus* C5 strain to the six strains using evolutionary genomic analysis (dN/dS)

We performed dN/dS analysis (branch-site model) to identify the evolutionarily selective genes in the *L. acidophilus* C5 strain. Considering the phylogenetic relationships among the seven *L. acidophilus* strains, we searched for the genes that could explain the specific characteristics of the *L. acidophilus* C5 strain. We identified 1,843 orthologous genes from the seven strains and measured the rate of evolution using the dN/dS analysis (). We identified 30 phylogenetically featured genes and the variations in their amino acid sequences (Supplementary Table S2). To determine the functions of the 30 evolutionarily selective genes, we mapped their sequences to the COG database (Fig. 6). We observed that apart from the "Function unknown" category, the largest proportion of core genes belonged to the "Carbohydrate transport and metabolism (G)" and "Transcription (K)" categories. The carbohydrate transport and metabolism (G) category included five genes-C5_1_01133 (glcU_1: glucose uptake protein GlcU), C5_1_00898 (fba: fructose-bisphosphate aldolase), C5_1_01372 (hypothetical protein), C5_1_01889 (ptsI: phosphoenolpyruvate-protein phosphotransferase), and C5_1_00253 (malL_2: Oligo-1,6-glucosidase).

DISCUSSION

This study is the first attempt to decipher the genetic features and evolutionary adaptations of *L. acidophilus* in the canine gut intestine. We performed whole-genome sequencing to construct the genome of the *L. acidophilus* C5 strain and compared the genomic information of *L. acidophilus* derived from a dog with six other strains isolated from diverse hosts. We determined the genetic basis for the characteristics of the *L. acidophilus* C5 strain that are likely related to the host. After the sequencing and assembly process, we were able to construct one large contig corresponding to the genome of the C5 strain, comparable in size (2.005 Mb) with the other six complete genomes of *L. acidophilus* (NCFM, LA14, FSI4, ATCC53544, LA1, and DSM20079) from NCBI (Table 1) [21–26]. In this study, complete genome sequences of other strains in the NCBI database were used for comparative genome analysis of C5, which were isolated from different hosts, and none were isolated from dogs as in previous studies [27,28]. In the annotation process, we found 2,082 genes (2,009 coding and 73 non-coding genes), which was slightly more than the number of genes

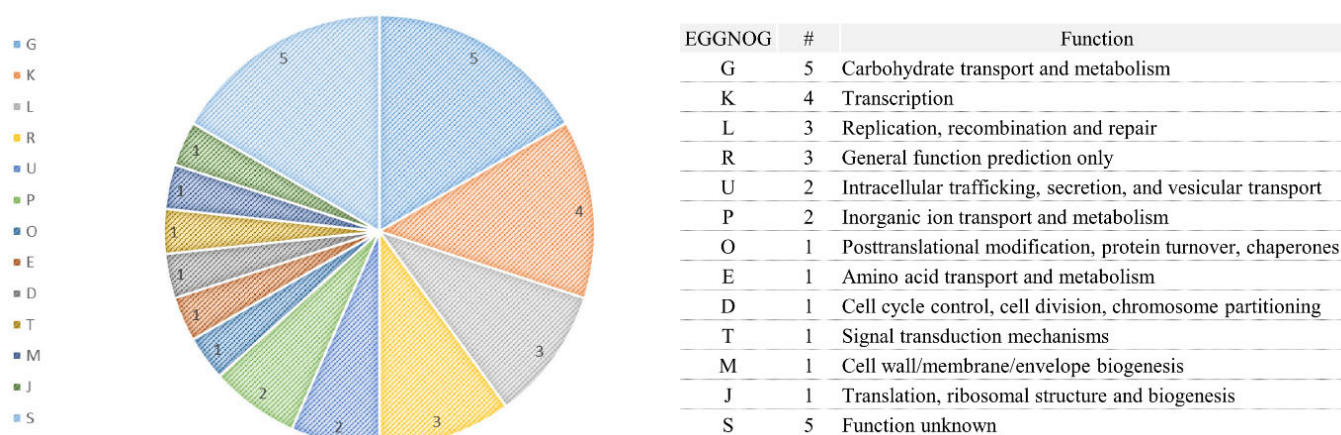| EGGNOG	#	Function
G	5	Carbohydrate transport and metabolism
K	4	Transcription
L	3	Replication, recombination and repair
R	3	General function prediction only
U	2	Intracellular trafficking, secretion, and vesicular transport
P	2	Inorganic ion transport and metabolism
O	1	Posttranslational modification, protein turnover, chaperones
E	1	Amino acid transport and metabolism
D	1	Cell cycle control, cell division, chromosome partitioning
T	1	Signal transduction mechanisms
M	1	Cell wall/membrane/envelope biogenesis
J	1	Translation, ribosomal structure and biogenesis
S	5	Function unknown

Fig. 6. Functional categorization using Eggnog annotation (COG database) of significant evolutionarily accelerated genes of *Lactobacillus acidophilus* C5 identified in the dN/dS analysis (the branch-site model). COG, clusters of orthologous groups; dN/dS, the ratio of non-synonymous to synonymous substitution.

found in the other six strains. Similarly, the genome size of *L. acidophilus* C5 was slightly larger than that of the other strains. In the functional analysis using the COG database, it was found that genes related to carbohydrate transport and metabolism (181 genes) comprised the largest part of the *L. acidophilus* C5 strain genome, apart from "General function prediction only" and "Function unknown" categories. In other strains too, the carbohydrate metabolism-related genes comprised a large portion of the genome (Fig. 4), but C5 was the only strain in which they were most abundant among the categories with distinct functions (except for "General function prediction only" and "Function unknown" categories).

In pan-genome analysis, a total of 1,726 core genes were detected in the seven *L. acidophilus* genomes isolated from four hosts, which mainly encoded essential proteins for metabolism (30.35%) (Fig. 4). Consistent with the findings of studies on other Lactobacillus strains, our findings suggested that core genes are indispensable, constitute the basic framework of the *L. acidophilus*, and play important roles in carbohydrate transport and metabolism (154 genes) and translation, ribosomal structure, and biogenesis (134 genes). Competitive retention of *L. acidophilus* in the intestinal tract is important for glycogen biosynthesis in carbohydrate metabolism, which demonstrates that the ability to synthesize intracellular glycogen contributes to gut fitness and indicates retention of probiotic microorganisms [29,30]. After pan-genome analysis, we constructed a phylogenetic tree based on the 1,726 core genes to evaluate the genetic relationships between the canine C5 strain and the six other strains of *L. acidophilus*. In this phylogenetic tree, we identified that C5 was distant from other strains but was close to two strains (DSM20079 and NCFM) from humans. This could be attributed to the fact that dogs are representative companion animals to humans and have shared their environments for a long time. Therefore, the genomic differences in *L. acidophilus* strains and similarities between the C5 and the two strains from humans might be associated with their colonized environments. Significantly, the strains closest to C5 were human-derived even if all human-derived strains were not close (such as LA1 and ATCC53544). Similar results were obtained in a comparative genome study of other Lactobacillus strains [27]. To better understand the characteristics of *L. acidophilus* C5, we investigated the unique genes of all *L. acidophilus* strains in this study. From the 2,254 gene clusters in pan-genome analysis, we found 328 unique genes, 245 of which were specific to the C5 strain. In the functional analysis based on the COG database, these C5-specific genes were mainly associated with "Replication, recombination, and repair (28 genes)" and "Carbohydrate transport and metabolism (24 genes)." Interestingly, core gene analysis revealed many common carbohydrate metabolism-related genes in the *L. acidophilus* strains, and it was confirmed that many carbohydrate metabolism-related genes in the C5 strain were not found in other strains. Moreover, we found that there were more carbohydrate metabolism-related genes in the C5 strain than in the other strains. Based on the core gene analysis and unique gene results, we inferred that C5 has distinct genomic features from other strains of *L. acidophilus*.

The similarity of several genomes within or between species is the basis of comparative genomics. If two species or strains have a recent common ancestor, the differences between the two genomes evolved after the common ancestral genome. The more closely related the two strains, the higher the similarities between their genomes [31]. When we performed comparative genomic analysis of the C5 strain isolated from dogs with other strains of *L. acidophilus*, we supposed that the genomic differences between the *L. acidophilus* strains were significant in the evolutionary process and could explain the adaption process of the C5 strain using the evolutionary statistical method (dN/dS) [32]. The identification of genetic loci undergoing adaptation is a central aim of evolutionary biology, and several statistical tests have been developed to quantify selection pressures acting on protein-coding regions. Among these methods, the dN/dS ratio is one of the most

widely used. It is simple and robust and can quantify selection pressures by comparing the rate of substitutions at silent sites (dS, which are presumed neutral) to the rate of substitutions at non-silent sites (dN, which possibly experience selection). The dN/dS ratio is used for distantly diverged sequences, so the differences among them represent substitutions that have been fixed along independent lineages [33,34]. We assumed that the selection signal by dN/dS indicated adaptation of each strain to its environment and used this measurement to understand the adaptation of the C5 strain to the canine environment [28, 35]. After identifying 1,843 orthologous genes in the seven genomes, we performed dN/dS estimation for each orthologous gene using PAML4 [20]. We determined 30 evolutionarily selective genes for the C5 strain. In these 30 genes, there were 300 C5-specific amino acid changes, out of which 141 were statistically significant. In the crcB gene, there were 30 C5-specific amino acid changes, the highest among the 30 selective genes, and 14 of all amino acid changes were statistically significant. This gene is important for preventing fluoride toxicity by reducing its concentration in the cell [36]. In the functional analysis based on the COG database, these evolutionarily selective genes were mainly associated with carbohydrate transport and metabolism (five genes) and transcription (four genes). Interestingly, among the genes containing significant evolutionary selection signals, most genes were carbohydrate metabolism-related genes. We hypothesized that the genes related to carbohydrate metabolism in strains isolated from dogs were closely related to the domestication of dogs. The domestication of dogs was an important milestone of human civilization [37]. In a previous study, whole-genome re-sequencing of dogs and wolves was performed to identify genomic regions potentially representing selection targets during dog domestication [38]. This study identified candidate mutations in key genes and provided functional support for increased starch digestion in dogs relative to wolves. This result indicates that adaptations allowed modern dog ancestors to survive on a diet rich in starch, a crucial step in the domestication of dogs. We wondered whether the evolution of *L. acidophilus* C5 from dogs has occurred. Reportedly, the dog and human gut microbiomes are similar in terms of gene content and response to diet [39]. As dogs and humans have shared a similar environment for a long time after domestication, if the dogs experience similar metagenome changes due to changes in diet, it was deduced that its microbiome adapts to a carbohydrate-rich diet. We inferred that more carbohydrate-related genes in the genome and genes with selection signals in *L. acidophilus* C5 were the result of evolution and domestication.

In summary our results indicated that the *L. acidophilus* C5 strain from a canine host had many genes related to carbohydrate metabolism, presumably due to domestication by humans. We reported the characteristics of the C5 strain from a canine using whole-genome sequencing data compared with other original isolates, providing a strong indication of the factors affecting its evolutionary history (evolution due to domestication of dogs by humans). We hope that our study contributes to the feasibility of using these strains as probiotics for dogs in the future.

SUPPLEMENTARY MATERIALS

Supplementary materials are only available online from: https://doi.org/10.5187/jast.2021.e126.

REFERENCES

1. Bâati L, Fabre-Gea C, Auriol D, Blanc PJ. Study of the cryotolerance of Lactobacillus acidophilus: effect of culture and freezing conditions on the viability and cellular protein levels. Int J Food Microbiol. 2000;59:241-7. https://doi.org/10.1016/S0168-1605(00)00361-5
2. Fijan S. Microorganisms with claimed probiotic properties: an overview of recent literature. Int

J Environ Res Public Health. 2014;11:4745-67. https://doi.org/10.3390/ijerph110504745

3. Ljungh A, Wadström T. Lactic acid bacteria as probiotics. Curr Issues Intest Microbiol. 2006;7:73-90.

4. Ghouri YA, Richards DM, Rahimi EF, Krill JT, Jelinek KA, DuPont AW. Systematic review of randomized controlled trials of probiotics, prebiotics, and synbiotics in inflammatory bowel disease. Clin Exp Gastroenterol. 2014;7:473-87. https://doi.org/10.2147/CEG.S27530

5. Marelli SP, Fusi E, Giardini A, Martino PA, Polli M, Bruni N, et al. Effects of probiotic Lactobacillus acidophilus D2/CSL (CECT 4529) on the nutritional and health status of boxer dogs. Vet Rec. 2020;187:e28. https://doi.org/10.1136/vr.105434

6. Barko PC, McMichael MA, Swanson KS, Williams DA. The gastrointestinal microbiome: a review. J Vet Intern Med. 2018;32:9-25. https://doi.org/10.1111/jvim.14875

7. Sommer F, Bäckhed F. The gut microbiota — masters of host development and physiology. Nat Rev Microbiol. 2013;11:227-38. https://doi.org/10.1038/nrmicro2974

8. Berger B, Pridmore RD, Barretto C, Delmas-Julien F, Schreiber K, Arigoni F, et al. Similarity and differences in the Lactobacillus acidophilus group identified by polyphasic analysis and comparative genomics. J Bacteriol. 2007:1311-21. https://doi.org/10.1128/JB.01393-06

9. Bull M, Plummer S, Marchesi J, Mahenthiralingam E. The life history of Lactobacillus acidophilus as a probiotic: a tale of revisionary taxonomy, misidentification and commercial success. FEMS Microbiol Lett. 2013;349:77-87. https://doi.org/10.1111/1574-6968.12293

10. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol. 2005;3:722-32. https://doi.org/10.1038/nrmicro1235

11. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol. 2017;13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595

12. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068-9. https://doi.org/10.1093/bioinformatics/btu153

13. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 2019;47:D309-14. https://doi.org/10.1093/nar/gky1085

14. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics. 2016;32:929-31. https://doi.org/10.1093/bioinformatics/btv681

15. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31:3691-3. https://doi.org/10.1093/bioinformatics/btv421

16. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:238. https://doi.org/10.1186/s13059-019-1832-y

17. Löytynoja A. Phylogeny-aware alignment with PRANK. In: Russell D, editor. Multiple sequence alignment methods. Totowa, NJ: Humana Press; 2014. p. 155-70.

18. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34:W609-12. https://doi.org/10.1093/nar/gkl315

19. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 2000;17:540-52. https://doi.org/10.1093/oxfordjournals.molbev.a026334

20. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol.

2007;24:1586-91. https://doi.org/10.1093/molbev/msm088

21. Pfeiler EA, Azcarate-Peril MA, Klaenhammer TR. Characterization of a novel bile-inducible operon encoding a two-component regulatory system in Lactobacillus acidophilus. J Bacteriol. 2007;189:4624-34. https://doi.org/10.1128/JB.00337-07

22. Stahl B, Barrangou R. Complete genome sequence of probiotic strain Lactobacillus acidophilus La-14. Genome Announc. 2013;1:e00376-13. https://doi.org/10.1128/genomeA.00376-13

23. Iartchouk O, Kozyavkin S, Karamychev V, Slesarev A. Complete genome sequence of Lactobacillus acidophilus FSI4, isolated from yogurt. Genome Announc. 2015;3:e00166-15. https://doi.org/10.1128/genomeA.00166-15

24. Dean SN, Vora GJ, Walper SA. Complete genome sequence of Lactobacillus acidophilus strain ATCC 53544. Genome Announc. 2017;5:e01138-17. https://doi.org/10.1128/genomeA.01138-17

25. Chung WH, Kang J, Lim MY, Lim T, Lim S, Roh SW, et al. Complete genome sequence and genomic characterization of Lactobacillus acidophilus LA1 (11869BP). Front Pharmacol. 2018;9:83. https://doi.org/10.3389/fphar.2018.00083

26. Falentin H, Cousin S, Clermont D, Creno S, Ma L, Chuat V, et al. Draft genome sequences of five strains of Lactobacillus acidophilus, strain CIP 76.13T, isolated from humans, strains CIRM-BIA 442 and CIRM-BIA 445, isolated from dairy products, and strains DSM 20242 and DSM 9126 of unknown origin. Genome Announc. 2013;1:e00658-13. https://doi.org/10.1128/genomeA.00658-13

27. Son S, Oh JD, Lee SH, Shin D, Kim Y. Comparative genomics of canine Lactobacillus reuteri reveals adaptation to a shared environment with humans. Genes Genomics. 2020;42:1107-16. https://doi.org/10.1007/s13258-020-00978-w

28. Heo J, Shin D, Chang SY, Bogere P, Park MR, Ryu S, et al. Comparative genome analysis and evaluation of probiotic characteristics of Lactobacillus plantarum strain JDFM LP11. Korean J Food Sci Anim Resour. 2018;38:878-8. https://doi.org/10.5851/kosfa.2018.e21

29. Goh YJ, Klaenhammer TR. Insights into glycogen metabolism in Lactobacillus acidophilus: impact on carbohydrate metabolism, stress tolerance and gut retention. Microb Cell Fact. 2014;13:94. https://doi.org/10.1186/s12934-014-0094-3

30. Goh YJ, Klaenhammer TR. A functional glycogen biosynthesis pathway in Lactobacillus acidophilus: expression and analysis of the glg operon. Mol Microbiol. 2013;89:1187-200. https://doi.org/10.1111/mmi.12338

31. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, et al. Comparative genomics of the lactic acid bacteria. Proc Natl Acad Sci USA. 2006;103:15611-6. https://doi.org/10.1073/pnas.0607117103

32. Kryazhimskiy S, Plotkin JB. The population genetics of dN/dS. PLOS Genet. 2008;4:e1000304. https://doi.org/10.1371/journal.pgen.1000304

33. Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature. 1977;267:275-6. https://doi.org/10.1038/267275a0

34. Yang Z, Bielawski JP. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 2000;15:496-503. https://doi.org/10.1016/S0169-5347(00)01994-7

35. Jang SY, Heo J, Park MR, Song MH, Kim JN, Jo SH, et al. Genome characteristics of Lactobacillus fermentum strain JDFM216 for application as probiotic bacteria. J Microbiol Biotechnol. 2017;27:1266-71. https://doi.org/10.4014/jmb.1703.03013

36. Baker JL, Sudarsan N, Weinberg Z, Roth A, Stockbridge RB, Breaker RR. Widespread genetic switches and toxicity resistance proteins for fluoride. Science. 2012;335:233-5. https://doi.org/10.1126/science.1215063

37. vonHoldt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature. 2010;464:898-902. https://doi.org/10.1038/nature08837

38. Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, et al. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. Nature. 2013;495:360-4. https://doi.org/10.1038/nature11837

39. Coelho LP, Kultima JR, Costea PI, Fournier C, Pan Y, Czarnecki-Maulden G, et al. Similarity of the dog and human gut microbiomes in gene content and response to diet. Microbiome. 2018;6:72. https://doi.org/10.1186/s40168-018-0450-3