Genome Characteristics of *Lactobacillus fermentum* Strain JDFM216 for Application as Probiotic Bacteria

Sung Yong Jang\(^1\†\), Jaeyoung Heo\(^2\†\), Mi Ri Park\(^1\), Min-Ho Song\(^3\), Jong Nam Kim\(^4\), Sung Ho Jo\(^5\), Do-Young Jeong\(^5\), Hak Kyo Lee\(^2\), Younghoon Kim\(^1\), and Sangnam Oh\(^1\)*

\(^1\)Department of Animal Science and Institute of Milk Genomics and \(^2\)Department of Animal Biotechnology, Chonbuk National University, Jeonju 54896, Republic of Korea

\(^3\)Department of Animal Science and Biotechnology, Chungnam National University, Daejeon 34134, Republic of Korea

\(^4\)Department of Beef & Dairy Science, Korea National College of Agriculture and Fisheries, Jeonju 54874, Republic of Korea

\(^5\)Microbial Institute for Fermentation Industry, Sunchang 56048, Republic of Korea

**Keywords:** *Lactobacillus fermentum*, genomic analysis, longevity, immune response

*Lactobacillus fermentum* strain JDFM216, isolated from a Korean infant feces sample, possesses the ability to enhance the longevity and immune response of a *Caenorhabditis elegans* host. To explore the characteristics of strain JDFM216 at the genetic level, we performed whole-genome sequencing using the PacBio system. The circular draft genome has a total length of 2,076,427 bp and a total of 2,682 encoding sequences were identified. Five phylogenetically featured genes possibly related to the longevity and immune response of the host were identified in *L. fermentum* strain JDFM216. These genes encode UDP-N-acetylglucosamine 1-carboxyvinyltransferase (E.C. 2.5.1.7), ErfK/YbiS/YcfS/YnhG family protein, site-specific recombinase XerD, homocysteine S-methyltransferase (E.C. 2.1.1.10), and aspartate-ammonia ligase (E.C. 6.3.1.1), which are involved in peptidoglycan synthesis and amino acid metabolism in the gut environment. Our findings on the genetic background of *L. fermentum* strain JDFM216 and its potential candidate genes for host longevity and immune response provide new insight for the application of this strain in the food industry as newly isolated functional probiotic.

Nowadays, it is an established fact that lactic acid bacteria (LAB) are generally recognized as safe microorganisms and responsible for the fermentation of food and feed [1, 2]. As probiotic bacteria, LAB benefit the host by improving the balance of gut microbial community, as well as by providing immunomodulatory effects [3, 4]. In recent years, interests in the host health-promoting effects of probiotics have been growing. As a result, various probiotic products have been developed to improve health [5]. In particular, numerous probiotic bacteria belonging to the *Lactobacillus* genus are commonly used as probiotics, and accepted as safe by the US Food and Drug Administration and the European Food Safety Authority [6]. Among them, *L. fermentum* is a heterofermentative natural inhabitant of the gastrointestinal tract that is often isolated from human biological samples (i.e., human breast milk and feces) [7]. This species has been previously reported for its different probiotic properties [6], whereas its genomic evidence is not yet available to support its functionality. In a preliminary study, we reported that *L. fermentum* strain JDFM216 isolated from a Korean infant feces sample enhances the lifespan and immune response of a *Caenorhabditis elegans* host. In the present study, we discovered the complete genome of *L. fermentum* strain JDFM216 that provides genomic evidence for the observed functions.

Genomic DNA of strain JDFM216 was obtained using an UltraClean Microbial DNA Isolation Kit (MoBio, USA) as described in the manufacturer’s protocol. The extracted DNA concentration and quality were measured using an Optizen NanoQ spectrophotometer (Optizen, Korea). An...
8–12 kb library was prepared following the Pacific Biosciences manual, and the sequencing procedure and filtration/de novo assembly of the raw sequence data was carried out on the PacBio RS II system (Pacific Biosciences, USA) using C4 chemistry on single-molecule real-time cells with a 120-min sequence capture protocol. The PacBio Hierarchical Genome Assembly Process and Quiver software package, respectively. Assembled contigs including a short length (<20,000 bp) and low coverage (<50×) were sorted for further experiment. Annotation of the JDFM216 genome was performed using the RAST annotation system [8], and COG annotation was conducted using previous methods as described [9]. For the comparative genome analysis, complete genomes or genome sequences of *L. fermentum* were searched on the NCBI database (https://www.ncbi.nlm.nih.gov/genome/genomes/711?). Average nucleotide identity (ANI) values were calculated for all 23 strains using JSpecies [10]. MESTORTH [11] and PRANK [12] methods were used for setting the orthologous gene for the eight complete genomes and for multiple sequence alignment of each orthologous gene, respectively. In addition, the construction of a phylogenetic tree with the neighbor-joining method was performed using MEGA6 [13]. Moreover, PAML4 analysis through the maximum-likelihood method [14] was used to estimate dS (synonymous substitution rate) and dN (nonsynonymous substitution rate) as well as phylogenetically featured genes investigated by the branch and branch-site model. To observe the stimulation of *pmk-1* (a key player for the p38 type mitogen-activated protein kinase pathway) in the nematode intestine, AY102 transgenic *Caenorhabditis elegans* (*P-vha-6::pmk-1::GFP + rol-6(su1006); pmk-1::GFP*) individuals were exposed on plates deposited with bacterial cells for 24 h. The nematodes were then mounted on glass slides with agarose pads and anesthetized with 10 mM NaN₃, and the transcriptional response of *pmk-1* was quickly visualized using fluorescence microscopy.

The *L. fermentum* strain JDFM216 genome possesses a single and circular type DNA chromosome of 2,076,427 bp.

**Fig. 1.** Genome map of *Lactobacillus fermentum* strain JDFM216. Marked characteristics are shown from outside to the center; CDS on forward strand, CDS on reverse strand, tRNA, rRNA, GC content, and GC skew. Functional genes are labeled around the outer circle as follows: phylogenetically featured genes in blue and genes related to antimicrobial activity in green. The scale is kilobase (kb) pair.
Table 1. Comparison of the chromosomal properties of *Lactobacillus fermentum* strains.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Sources</th>
<th>3872</th>
<th>CECT5716</th>
<th>F-6</th>
<th>IFO3956</th>
<th>JDFM216</th>
<th>NCC2970</th>
<th>SNUV175</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Plasmids</td>
<td>1</td>
<td>-</td>
<td>NC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Genome size (Mb)</td>
<td>2.33049</td>
<td>2.10045</td>
<td>2.06462</td>
<td>2.09869</td>
<td>2.07643</td>
<td>1.94987</td>
<td>2.27233</td>
<td></td>
</tr>
<tr>
<td>GC content (%)</td>
<td>50.6</td>
<td>51.5</td>
<td>51.7</td>
<td>51.5</td>
<td>51.5</td>
<td>52.2</td>
<td>51.1</td>
<td></td>
</tr>
<tr>
<td>Open reading frames</td>
<td>2,328</td>
<td>2,097</td>
<td>2,041</td>
<td>2,102</td>
<td>2,682</td>
<td>1,912</td>
<td>2,312</td>
<td></td>
</tr>
<tr>
<td>Annotated genes</td>
<td>1,863</td>
<td>1,717</td>
<td>1,717</td>
<td>1,716</td>
<td>2,085</td>
<td>1,614</td>
<td>1,863</td>
<td></td>
</tr>
<tr>
<td>Hypothetical and unknown proteins</td>
<td>465</td>
<td>380</td>
<td>324</td>
<td>386</td>
<td>597</td>
<td>298</td>
<td>449</td>
<td></td>
</tr>
<tr>
<td>tRNA</td>
<td>58</td>
<td>57</td>
<td>58</td>
<td>57</td>
<td>58</td>
<td>37</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>rRNA</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>14</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Average nucleotide identity (%)</td>
<td>97.9</td>
<td>97.8</td>
<td>97.7</td>
<td>97.9</td>
<td>100.0</td>
<td>97.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NC: not confirmed.

Fig. 2. Functional categorization of all predicted open reading frames in the *Lactobacillus fermentum* strain JDFM216 genome based on the (A) COG and (B) SEED databases.
with a GC content of 51.48%. The JDFM216 genome has a total of 2,682 open reading frames (ORFs) as well as 59 tRNAs and 15 rRNAs (Fig. 1 and Table 1). Importantly, the ORFs of JDFM216 were larger in number than those of the other L. fermentum strains, even though the genome size was similar. Generally, a number of plasmids are different in strains. Based on our bioinformatics analysis, no plasmid could be detected in JDFM216, similar to other strains, including CECT 5716, F-6, IFO 3956, and NCC 2970, but not 3872 and SNUV175. Among the identified ORFs, 2,085 genes (77.7%) were predicted as functional genes and 597 (32.3%) were unknown or hypothetical genes. As shown in Fig. 2, the predicted ORFs were grouped by COG functional and SEED subsystem categorizations. First, COG functional categorization showed that 1,075 ORFs (51.56% of the COG assigned ORFs) belonged to five major COG functional categories, including amino acid transport and metabolism, carbohydrate transport and metabolism, translation, ribosomal structure and biogenesis, recombination, replication, repair, and general function prediction (Fig. 2A). In addition, SEED categorization mainly resulted in ORFs responsible for RNA metabolism, amino acids and derivatives, and DNA metabolism as well as carbohydrate, vitamins, prosthetic groups, and cofactor (54.02% in 2,085 ORFs).

Next, we constructed two ANI trees and one phylogenetic tree for a comparative tree analysis of JDFM216 and its same species. The two ANI trees were constructed by 15 available genome sequences and seven available complete genome sequences in the NCBI database, respectively (Fig. 3). Six strains, namely CECT_5716, IFO_3956, F-6, NCC2970, SNUV175, and 3872, were grouped together with JDFM216. Unexpectedly, the topological patterns of the phylogenetic tree did not completely correspond to that of the ANI tree. In the phylogenetic tree, however, JDFM216 clustered with SNUV175. These results implied that the sequences of genome possibly are dependent on strain or origin, even though they involved the same species. Additionally, dN/dS analysis was performed to identify the phylogenetically featured genes in the JDFM216 strain (Table 2). Unfortunately, there was no identification of featured genes in the branch model (data not shown). Based on the dN/dS analysis, five phylogenetically featured genes were revealed in this study (Table 2); UDP-N-acetylglucosamine 1-carboxyvinyltransferase (E.C. 2.5.1.7),

![Fig. 3. Comparative tree analysis.](image)

(A) ANI tree analysis of 15 available genome sequences in the Lactobacillus fermentum strain using JSpecies. (B) ANI tree analysis of JDFM216 with six available complete genome sequences of L. fermentum. (C) Phylogenetic tree analysis of JDFM216 with six available complete genome sequences.
ErfK/YbiS/YcfS/YnhG family protein, site-specific recombinase XerD, homocysteine S-methyltransferase (EC 2.1.1.10), and aspartate-ammonia ligase (EC 6.3.1.1). Among them, UDP-N-acetylglucosamine 1-carboxyvinyltransferase and ErfK/YbiS/YcfS/YnhG family of proteins are involved in peptidoglycan biosynthesis, but their biological roles are not clear. Peptidoglycan (PG) is a major component in bacterial cell wall and acts as microbe-associated molecular patterns in bacteria-host interaction [15]. Recently, it has been accepted that PG generally influences the host’s immunity as well as the host’s physiology, including mammalian aging [16, 17]. Previously, we indicated that the pmk-1 pathway is a key player of C. elegans immune response and the pmk-1 signaling pathway is strongly involved in enhanced immunity under intake of probiotic bacteria [18]; hence, we also determined the stimulation of pmk-1 via transgenic worms harboring the pmk-1::GFP reporter system with L. fermentum strain JDFM216. Expectedly, we observed dramatic stimulation of pmk-1::GFP in the presence of conditioning with L. fermentum strain JDFM216, but not with L. plantarum strain JDFM675 (Fig. 4). Therefore, we considered that PG-associated factors encoded by featured genes in L. fermentum strain JDFM216 may stimulate intestinal epithelial cells and also influence the immune response and longevity of the host.

In addition, homocysteine S-methyltransferase, a key player in methionine metabolism [19], was identified in the present study. Notably, administration of methionine in diets resulted in lifespan extension in rats and mice [20], as well as in Drosophila melanogaster [21] and C. elegans via gut microbiota [22]. Taken together, our results suggest that the featured genes, including those encoding for cell wall components and amino acid metabolism, might improve the lifespan and influence the immune response of the host.
In conclusion, we showed that potential candidate genes of *L. fermentum* strain JDFM216 may be involved in enhancement of the longevity and host immune response via phylogenetically featured genes related to PG synthesis and amino acid metabolism, including methionine in the gut environments. As indicated, whole-genome sequencing results showed that JDFM216 is different from the complete genome sequences of the other strains of this species (<97% of ANI compared with others; Table 2). Therefore, the strain-to-strain variation in *L. fermentum* may reflect the probiotic functionality of this microorganism on the host. Ongoing study is evaluating these phylogenetically featured genes to clarify the health-promoting mechanisms of *L. fermentum* strain JDFM216 in the gastrointestinal tract.

**Acknowledgments**

This research was supported by the Ministry of Trade, Industry & Energy (MOTIE), Korea Institute for Advancement of Technology (KIAT) and Establishment of Infrastructure for Industrialization of Korean Useful Microbes (R0004073), and by a grant from the Next-Generation BioGreen 21 Program (PJ01104401), Rural Development Administration, Republic of Korea.

**References**


July 2017 | Vol. 27 | No. 7