**ARTICLE INFORMATION**

<table>
<thead>
<tr>
<th><strong>ARTICLE INFORMATION</strong></th>
<th><strong>Fill in information in each box below</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Article Type</td>
<td>Research article</td>
</tr>
<tr>
<td>Article Title (within 20 words without abbreviations)</td>
<td>Complete genome sequence of <em>Lactilactobacillus curvatus</em> CACC879 and its functional probiotic properties</td>
</tr>
<tr>
<td>Running Title (within 10 words)</td>
<td>Complete genome sequence of <em>Lactilactobacillus curvatus</em> CACC879</td>
</tr>
<tr>
<td>Author</td>
<td>Soyeon Park¹, Seoyun Son¹, Mi Ae Park¹, Dae-Hyuk Kim¹,² and Yangseon Kim¹</td>
</tr>
<tr>
<td>Affiliation</td>
<td>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</td>
</tr>
<tr>
<td>ORCID (for more information, please visit <a href="https://orcid.org">https://orcid.org</a>)</td>
<td>Soyeon Park (<a href="http://orcid.org/0000-0003-3788-5415">http://orcid.org/0000-0003-3788-5415</a>) Seoyun Son (<a href="http://orcid.org/0000-0002-4753-1955">http://orcid.org/0000-0002-4753-1955</a>) Mi Ae Park (<a href="http://orcid.org/0000-0002-7601-5976">http://orcid.org/0000-0002-7601-5976</a>) Dae-Hyuk Kim (<a href="http://orcid.org/0000-0002-9948-5313">http://orcid.org/0000-0002-9948-5313</a>) Yangseon Kim (<a href="http://orcid.org/0000-0002-8285-3407">http://orcid.org/0000-0002-8285-3407</a>)</td>
</tr>
<tr>
<td>Competing interests</td>
<td>No potential conflict of interest relevant to this article was reported</td>
</tr>
<tr>
<td>Funding sources</td>
<td>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 2022 Technology commercialization support project Ministry of Agriculture, Food and Rural Affairs, Republic of Korea (Project No. 122037–02)</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Availability of data and material</td>
<td>Upon reasonable request, the datasets of this study can be available from the corresponding author</td>
</tr>
<tr>
<td>Authors’ contributions</td>
<td>Conceptualization: Kim Y. Data curation: Park S, Park MA. Formal analysis: Park S. Methodology: Park S, Son S. Software: Park S. Validation: Park S, Kim DH, Son S. Investigation: Kim Y. Writing - original draft: Park S. Writing - review &amp; editing: Park S, Son S, Park MA, Kim DH, Kim Y.</td>
</tr>
<tr>
<td>Ethics approval and consent to participate</td>
<td>This article does not require IRB/IACUC approval because there are no human and animal participants</td>
</tr>
<tr>
<td>For the corresponding author (responsible for correspondence, proofreading, and reprints)</td>
<td>Fill in information in each box below</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>First name, middle initial, last name</td>
<td>Yangseon Kim</td>
</tr>
<tr>
<td>Email address – this is where your proofs will be sent</td>
<td><a href="mailto:yangseon@cialm.or.kr">yangseon@cialm.or.kr</a></td>
</tr>
<tr>
<td>Secondary Email address</td>
<td><a href="mailto:nimitzdr93@gmail.com">nimitzdr93@gmail.com</a></td>
</tr>
<tr>
<td>Address</td>
<td>241, Cheomdangwahagro, Jeongeup-si, Jeonbuk, Korea</td>
</tr>
<tr>
<td>Cell phone number</td>
<td>82-10-5680-4340</td>
</tr>
<tr>
<td>Office phone number</td>
<td>82-63-536-6712</td>
</tr>
<tr>
<td>Fax number</td>
<td>82-63-536-6003</td>
</tr>
</tbody>
</table>
Abstract

_Latilactobacillus curvatus_ CACC879 originated from swine feces in Korea, and its probiotic properties have been analyzed. The complete genome of strain CACC879 contained one chromosome 1,398,247 bp in length and three circular plasmids, namely, pCACC879-1 (591,981 bp), pCACC879-2 (14,542 bp), and pCACC879-3 (45,393 bp). The complete genome encodes a total of 2,077 genes, including 25 rRNA genes and 90 tRNA genes. In addition, probiotic stability-genes acid/bile related to salts tolerance, the biosynthesis of cobalamin (vitamin B12), riboflavin (vitamin B2), and CRISPR/Cas9 were found in the whole genomes. Remarkably, _L. curvatus_ CACC879 contained the antioxidant-related (peroxiredoxin) and bacteriocin-related genes (_lysM and blpA_). Overall, these results demonstrate that _L. curvatus_ CACC879 is a functional probiotic candidate for animal industry applications.

Keywords: _Latilactobacillus curvatus_, swine, probiotics, PacBio, genome sequence
Lactic acid bacteria, such as *Latilactobacillus*, are useful microbes that produce healthy metabolites, including bacteriocins and organic acids (such as lactic acid), that can regulate the gut microbiome balance [1]. Lactic acid bacteria also confer health benefits via diverse mechanisms, such as acid and bile tolerance, epithelial cell adherence, intestinal barrier buildup, and immune system modulation [2]. *Latilactobacillus curvatus* is a potential probiotic strain that produces various bacteriocins and metabolites and exhibits immunomodulatory activity [3, 4]. In this study, the genome of *L. curvatus* CACC879 was sequenced and fully assembled to elucidate the genetic factors associated with its probiotic characteristics.

*L. curvatus* CACC879 was isolated from swine feces in Korea, and the isolate was cultured in De Man, Rogosa, and Sharpe (MRS) medium for 18 h at 37 °C. The genomic DNA of *L. curvatus* CACC879 was extracted and purified using the DNeasy UltraClean kit (Qiagen, Hilden, Germany) and sequenced using the PacBio Sequel (Pacific Biosciences, Menlo Park, CA, USA) sequencing platform. De novo assembly was performed using PacBio SMRT analysis software (version 2.3.0; Pacific Biosciences) [5]. The EggNOG 5.0 database (http://eggnog5.embl.de) was used to classify all genes into clusters of ortholog gene (COG)/non-supervised orthologous group (NOG) categories. Functional annotations of the predicted coding sequences (CDSs) were compared with the Swiss-Prot and Kyoto Encyclopedia of Genes and Genomes (KEGG) [6]. The genome sequence of CACC879 was compared with other reference strains by Orthologous average nucleotide identity (OrthoANI; http://www.ezbiocloud.net/tools/orthoani) [7].

The whole genome of strain CACC879 consisted of one circular chromosome 1,398,247 bp in length (41.9% GC) along with three plasmids designated pCACC879-1 (591,981 bp, 42.2% GC), pCACC879-2 (14,542 bp, 45.2% GC), and pCACC879-3 (45,393 bp, 41.2% GC) (Table 1 and Fig. 1A). The genome of strain CACC879 contains 2,077 CDSs and 115 non-coding genes.
(25 rRNA and 90 tRNA genes) (Table 1). In addition, a total of 1,874 proteins (90.2%) were matched and classified into 19 COG functional categories (Fig. 1B). The most abundant COG categories were associated with replication, recombination, and repair (12.7%); translation, ribosomal structure, and biogenesis (7.8%); transcription (7.6%), carbohydrate transport, and metabolism (7.5%); and cell wall/membrane/envelope biogenesis (5.7%), excluding those with unknown function (29.7%). Compared with the genome sequence of reference strains, the genome of strain CACC879 was the most similar to that of the reference strains *L. curvatus* DSM 20019 (99.4%) and Wikim38 (99.0%) (Fig. 1C). The CACC879 strain showed common probiotic properties including the CRISPR-associated endonuclease (Cas9) for antiviral-related mechanisms and the biosynthesis of vitamin B groups (*ribF* and * pduO*), bacteriocin (*lysM*), and antioxidant (*tpx*), compared to the reference strains [8-10]. Additionally, we confirmed that strain CACC879 harbors genes associated with common probiotic properties, including acid tolerance (*clpB* and *grpE*), lactate synthesis (*ldh* and L-lactate dehydrogenase), and cell adhesion (*sotA*) (Table 2). Interestingly, the CACC879 genome contained the *dltB* and *dltD* genes associated with the modulation of the host immune response, but the reference strains did not. These findings will serve as a reference for further studies on *L. curvatus* and provide a scientific basis for functional probiotic development.
Nucleotide sequence accession number(s)

The whole-genome sequence of *L. curvatus* strain CACC879 (KACC 92511) has been deposited in GenBank under accession numbers CP117683 (chromosome) and CP117684 to CP117686 (plasmids). The BioProject and BioSample accession numbers are PRJNA932593 and SAMN33197937.
Acknowledgement

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 2022 Technology commercialization support project Ministry of Agriculture, Food and Rural Affairs, Republic of Korea (Project No. 122037–02).
References


Table 1. Genome features of *Lactilactobacillus curvatus* CACC879

<table>
<thead>
<tr>
<th>Properties</th>
<th>Chromosome CACC879</th>
<th>Plasmids pCACC879-1</th>
<th>pCACC879-2</th>
<th>pCACC879-3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (bp)</td>
<td>1,398,247</td>
<td>591,981</td>
<td>14,542</td>
<td>45,393</td>
<td>2,050,163</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>41.9</td>
<td>42.2</td>
<td>45.2</td>
<td>41.2</td>
<td>42.0</td>
</tr>
<tr>
<td>CDSs</td>
<td>1,421</td>
<td>596</td>
<td>11</td>
<td>49</td>
<td>2,077</td>
</tr>
<tr>
<td>tRNA</td>
<td>35</td>
<td>27</td>
<td>28</td>
<td>–</td>
<td>90</td>
</tr>
<tr>
<td>rRNA</td>
<td>14</td>
<td>8</td>
<td>3</td>
<td>–</td>
<td>25</td>
</tr>
<tr>
<td>CRISPR regions</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
</tbody>
</table>

GC, guanine and cytosine; CDSs, coding DNA sequences; tRNA, transfer RNA; rRNA, ribosomal RNA; CRISPR, clustered regularly interspaced short palindromic repeats.
Table 2. Predicted CDSs involved in *Latilactobacillus curvatus* CACC879 probiotic potency

<table>
<thead>
<tr>
<th>Predicted function</th>
<th>Gene</th>
<th>Start</th>
<th>End</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRISPR-associated endonuclease</td>
<td>Cas9</td>
<td>31,830</td>
<td>35,723</td>
<td>3,894</td>
</tr>
<tr>
<td>Antimicrobial activity-related</td>
<td>lysM</td>
<td>1,089,047</td>
<td>1,090,990</td>
<td>1,944</td>
</tr>
<tr>
<td>Bacteriocin (Class II)-related</td>
<td>blpA</td>
<td>c557,488</td>
<td>c558,636</td>
<td>1,149</td>
</tr>
<tr>
<td>Lactate synthesis</td>
<td>ldh</td>
<td>c903,407</td>
<td>c904,384</td>
<td>978</td>
</tr>
<tr>
<td><strong>Acid tolerance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chaperone protein ClpB</td>
<td>clpB</td>
<td>706,508</td>
<td>707,140</td>
<td>633</td>
</tr>
<tr>
<td>Chaperone protein GrpE</td>
<td>grpE</td>
<td>1,335,554</td>
<td>1,336,165</td>
<td>612</td>
</tr>
<tr>
<td>CIC family H(+)/Cl(−) exchange transporter</td>
<td>eriC</td>
<td>1,148,397</td>
<td>1,149,971</td>
<td>1,575</td>
</tr>
<tr>
<td>Sodium hydrogen exchanger family protein</td>
<td>nhaP</td>
<td>1,219,349</td>
<td>1,221,493</td>
<td>2145</td>
</tr>
<tr>
<td>F0F1 ATP synthase subunit A</td>
<td>atpB</td>
<td>17,075</td>
<td>17,788</td>
<td>714</td>
</tr>
<tr>
<td>F0F1 ATP synthase subunit B</td>
<td>atpF</td>
<td>18,084</td>
<td>18,605</td>
<td>522</td>
</tr>
<tr>
<td>F0F1 ATP synthase subunit C</td>
<td>atpE</td>
<td>17,807</td>
<td>18,019</td>
<td>213</td>
</tr>
<tr>
<td>F0F1 ATP synthase subunit delta</td>
<td>atpH</td>
<td>18,592</td>
<td>19,134</td>
<td>543</td>
</tr>
<tr>
<td><strong>Bile salts tolerance</strong></td>
<td>cbh</td>
<td>c246,195</td>
<td>C246,716</td>
<td>522</td>
</tr>
<tr>
<td><strong>Cell adhesion</strong></td>
<td>satA</td>
<td>73,428</td>
<td>74,090</td>
<td>663</td>
</tr>
<tr>
<td><strong>Stress response or protection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chaperone protein DnaK</td>
<td>dnaK</td>
<td>1,336,202</td>
<td>1,338,037</td>
<td>1,836</td>
</tr>
<tr>
<td>Chaperone protein DnaJ</td>
<td>dnaJ</td>
<td>1,338,161</td>
<td>1,339,309</td>
<td>1,149</td>
</tr>
<tr>
<td>Triose-phosphate isomerase</td>
<td>tpiA</td>
<td>c532,378</td>
<td>c533,133</td>
<td>756</td>
</tr>
<tr>
<td><strong>Biosynthesis of vitamin B groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riboflavin (B2)</td>
<td>ribF</td>
<td>1,332,094</td>
<td>1,333,047</td>
<td>954</td>
</tr>
<tr>
<td>Cobalamin (B12)</td>
<td>ribT</td>
<td>c271,721</td>
<td>c272,089</td>
<td>369</td>
</tr>
<tr>
<td><strong>Modulation of Immune response</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-alanyl-lipoteichoic acid biosynthesis proteins</td>
<td>dltB</td>
<td>c220,566</td>
<td>c221,774</td>
<td>1,209</td>
</tr>
<tr>
<td></td>
<td>dltD</td>
<td>c219,030</td>
<td>c220,298</td>
<td>1,269</td>
</tr>
<tr>
<td>Antioxidant (peroxiredoxin)</td>
<td>tpx</td>
<td>565,480</td>
<td>565,974</td>
<td>495</td>
</tr>
<tr>
<td></td>
<td>tpxA</td>
<td>780,278</td>
<td>780,601</td>
<td>324</td>
</tr>
</tbody>
</table>

CDSs, coding sequences; CRISPR, clustered regularly interspaced short palindromic repeats
Figure 1. Genome features of *Lactilactobacillus curvatus* CACC879. (A) Circular genome mapping of *L. curvatus* CACC879. (B) Functional classification of clusters of orthologous groups (COG). (C) Orthologous average nucleotide identity (OrthoANI) values of *L. curvatus* CACC879 compared to other reference strains.