Complete genome sequence of *Leuconostoc suionicum* DSM 20241<sup>T</sup> provides insights into its functional and metabolic features

Byung Hee Chun<sup>††</sup>, Se Hee Lee<sup>††</sup>, Hye Hee Jeon<sup>†</sup>, Dong-Woon Kim<sup>‡</sup> and Che Ok Jeon<sup>†*</sup>

Abstract

The genome of *Leuconostoc suionicum* DSM 20241<sup>T</sup> (=ATCC 9135<sup>T</sup> = LMG 8159<sup>T</sup> = NCIMB 6992<sup>T</sup>) was completely sequenced and its fermentative metabolic pathways were reconstructed to investigate the fermentative properties and metabolites of strain DSM 20241<sup>T</sup> during fermentation. The genome of *L. suionicum* DSM 20241<sup>T</sup> consists of a circular chromosome (2026.8 Kb) and a circular plasmid (21.9 Kb) with 37.58% G + C content, encoding 997 proteins, 12 rRNAs, and 72 tRNAs. Analysis of the metabolic pathways of *L. suionicum* DSM 20241<sup>T</sup> revealed that strain DSM 20241<sup>T</sup> performs heterolactic acid fermentation and can metabolize diverse organic compounds including glucose, fructose, galactose, cellobiose, mannose, sucrose, trehalose, arabinit, salcin, xylose, arabinose and ribose.

Keywords: *Leuconostoc suionicum*, Complete genome, Lactic acid bacteria, KEGG, Fermentative metabolic pathway

Introduction

The genus *Leuconostoc* comprises Gram-positive, facultatively anaerobic, intrinsically vancomycin-resistant, catalase-negative, spherical heterofermentative lactic acid bacteria which are involved in the fermentation of plant materials (such as kimchi), dairy products, meats, vegetable sausages and beverages [1–7]. Strain DSM 20241<sup>T</sup> (=ATCC 9135<sup>T</sup> = LMG 8159<sup>T</sup> = NCIMB 6992<sup>T</sup>) of the genus *Leuconostoc* was isolated in Sweden in 1972. It was originally classified as a subspecies of *L. mesenteroides*, but was recently reclassified as a novel species – *L. suionicum* – based on its whole genome sequence [4]. Here, we present the taxonomic and genomic features of *L. suionicum* DSM 20241<sup>T</sup>. In addition, we investigated the metabolic properties of *L. suionicum* DSM 20241<sup>T</sup> and reconstructed the metabolic pathways of organic compounds to estimate the fermentative metabolites in *L. suionicum* DSM 20241<sup>T</sup>.

Organism information

Classification and features

*L. suionicum* DSM 20241<sup>T</sup> belongs to the family *Leuconostocaceae*, order *Lactobacillales*, class *Bacilli* and phylum *Firmicutes*. Strain DSM 20241<sup>T</sup> is a Gram-positive, facultatively anaerobic, non-motile, non-sporulating, catalase-negative coccus, with a diameter of 0.5–0.7 μm (Fig. 1). It can be grown in MRS broth at 10–40 °C, with an optimal growth temperature of 30 °C [4]. Strain DSM 20241<sup>T</sup> ferments a wide variety of carbon sources including D-glucose, arbutin, melibiose, sucrose, turanose, N-acetylg glucosamine, cellobiose, galactose, gentiobiose, amygdalin, L-arabinose, esculin, ferric citrate, D-fructose, D-mannose, lactose, maltose, methyl α-D-glucopyranoside, salicin, trehalose, D-xylose, potassium 5-ketogluconate, mannitol and ribose to produce gas and acids (Table 1); however, it does not ferment glycerol, erythritol, D-arabinose, L-xylose, D-adonitol, methyl β-D-xylopyranoside, L-sorbose,
methyl α-D-mannopyranoside, L-rhamnose, dulcitol, inositol, D-sorbitol, inulin, D-melezitose, starch, glycogen, xylitol, D-lyxose, D-tagatose, fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate or raffinose [4, 8].

Phylogenetic analysis using the 16S rRNA gene sequences with validated type strains showed that L. suionicum DSM 20241 T is most closely related to the subspecies of the species L. mesenteroides: L. mesenteroides subsp. mesenteroides, L. mesenteroides subsp. jonggajibkimchii, L. mesenteroides subsp. cremoris, and L. mesenteroides subsp. dextranicum with very high 16S rRNA gene sequence similarities (>99.73%; Fig. 2).  

Genome sequencing information

Genome project history
L. suionicum DSM 20241 T was selected owing to its taxonomic significance for the species L. mesenteroides and was obtained from the German Collection of Microorganisms and Cell Cultures. The complete sequences of the chromosome and plasmid of strain DSM 20241 T were deposited in GenBank with the accession numbers CP015247–48. The project information and its association with MIGS version 2.0 [9] are summarized in Table 2.

Growth conditions and genomic DNA preparation
L. suionicum DSM 20241 T was cultured in MRS broth (BD Biosciences, CA, USA) at 30 °C for 24 h until the early stationary phase. Genomic DNA was extracted according to a standard phenol-chloroform extraction and ethanol precipitation procedure [10]. DNA quality (OD260/OD280 > 1.8) and concentration were measured using a NanoDrop ND-1000 spectrophotometer (Synergy Mx, Biotek, VT, USA).

Table 1 Classification and general features of Leuconostoc suionicum DSM 20241 T according to MIGS recommendations [9]

| MIGS ID | Property | Term | Evidence codea |
|---------|----------|------|---------------|
| Classification | Domain | Bacteria | TAS [31] |
| | Phylum | Firmicutes | TAS [32, 33] |
| | Class | Bacilli | TAS [34] |
| | Order | Lactobacillales | TAS [35] |
| | Family | Leuconostocaceae | TAS [35] |
| | Genus | Leuconostoc | TAS [36–38] |
| | Species | Leuconostoc suionicum | TAS [4] |
| | Type strain | DSM 20241 T | TAS [4] |
| Gram stain | | Positive | TAS [8] |
| Cell shape | | Coccus | TAS [8] |
| Motility | | Non-motile | NAS |
| Sporulation | | Non-sporulating | TAS [8] |
| Temperature range | | 10–40 °C | TAS [4] |
| Optimum temperature | | 30 °C | TAS [4, 8] |
| pH range; Optimun | | Not reported | NAS |
| Carbon source | l-arabinose, ribose, D-xylose, galactose, glucose, fructose, mannose, methyl α-D-glucopyranoside, N-acetylglucosamine, amygdalin, arbutin, aesculin, salicin, cellobiose, maltose, melibiose, sucrose, trehalose, gentiobiose and turanose | TAS [4, 8] |

aEvidence codes - IDA Inferred from Direct Assay, TAS Traceable Author Statement (i.e., a direct report exists in the literature), NAS Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [cite this reference]
Genome sequencing and assembly

The genome of strain DSM 20241\textsuperscript{T} was sequenced using PacBio RS SMRT technology based on a 10-kb SMRT-bell library at Macrogen (Seoul, Korea) as previously described [10]; 138,738 high-quality reads were generated, with an average length of 7656 bp. De novo assembly of sequencing reads derived from PacBio SMRT sequencing was performed using the hierarchical genome assembly process (HGAP; ver. 3.0) [11], which yielded a circular chromosome (2,026,850 bp) and a circular plasmid (21,983 bp) (Fig. 3).

Genome annotation

Automated genome annotation of strain DSM 20241\textsuperscript{T} was performed using Prodigal as part of the Joint Genome Institute's microbial genome annotation pipeline [12]. In addition, predicted coding sequences were functionally annotated using the NCBI non-redundant database, UniProt, TIGR-Fam, Pfam, PRIAM, Kyoto Encyclopedia of Genes and Genomes, Clusters of Orthologous Groups, and InterPro. Structural RNA genes were identified by using HMMER 3.0rc1 (rRNAs) [13] and tRNAscan-SE 1.23 (tRNAs) [14]. Other non-coding genes were searched using INFERNAL 1.0.2 [15]. Additional annotation was performed within the

---

**Table 2** Genome sequencing project information for *Leuconostoc suionicum* DSM 20241\textsuperscript{T}

| MIGS ID | Property                     | Term                                           |
|--------|------------------------------|------------------------------------------------|
| MIGS 31| Finishing quality            | Complete                                       |
| MIGS-28| Libraries used               | PacBio 10-kb SMRT-bell library                 |
| MIGS 29| Sequencing platforms         | PacBio RS SMRT                                 |
| MIGS 31.2| Fold coverage               | 50 ×                                           |
| MIGS 30| Assemblers                   | RS_HGAP Assembly.3                            |
| MIGS 32| Gene calling method          | NCBI Prokaryotic Genome, Annotation Pipeline   |
| Locus Tag         |                              | A6B45                                          |
| GenBank ID        |                              | CP015247-CP015248                              |
| GenBank Date of Release |                          | 14-APR-2017                                    |
| GOLD ID           |                              | Ga0151201                                      |
| BIOPROJECT        |                              | PRJNA318320                                    |
| MIGS 13| Source Material Identifier   | DSM 20241\textsuperscript{T}/ATCC 9135\textsuperscript{T}/LMG 8159\textsuperscript{T}/NCIMB 6992\textsuperscript{T} |
| Project relevance |                              | Taxonomy, industry, fermentation               |

---

**Fig. 2** Neighbor-joining tree based on the 16S rRNA gene sequences showing the phylogenetic relationships between *Leuconostoc suionicum* DSM 20241\textsuperscript{T} (highlighted in bold) and closely related *Leuconostoc* species. The sequences were aligned using the fast secondary-structure aware Infernal aligner available from the Ribosomal Database Project [28] and the tree was constructed based on the neighbor-joining algorithm using PHYLIP software (ver. 3.68) [29]. Bootstrap values of over 70% are shown on the nodes as percentages of 1000 replicates. Weissella viridescens 1536\textsuperscript{T} (AB023236) was used as an outgroup (not shown). Bar indicates 0.01 changes per nucleotide position.
Integrated Microbial Genomes—Expert Review platform [16].

**Genome properties**
The complete genome of *L. suionicum* strain DSM 20241<sup>T</sup> consists of a circular chromosome (2,026,850 bp) and a circular plasmid (21,983 bp) with 37.6% and 37.0% G + C contents, respectively (Table 3). The genome contains 1997 protein coding genes and 93 RNA genes (72 tRNAs, 12 rRNAs and 9 other RNAs; Table 4). Additional genome statistics and the distribution of the genes into COG functional categories are presented in Tables 4 and 5, respectively.

**Insights from the genome sequence**

**KEGG metabolic and regulatory pathways**
The KEGG metabolic pathways of *L. suionicum* DSM 20241<sup>T</sup> show that strain DSM 20241<sup>T</sup> displays typical heterolactic acid fermentative capabilities, performing pentose phosphate metabolism, fructose and mannose metabolism, galactose metabolism, sucrose metabolism and pyruvate metabolism without the complete tricarboxylic acid cycle (Fig. 4a, see Additional file 1: Table S1) [17–19]. In addition, *L. suionicum* DSM 20241<sup>T</sup> harbors genes related to riboflavin metabolism, fatty acid biosynthesis, purine and pyrimidine metabolism and amino acid biosynthesis (Fig. 4a). The regulatory pathways of strain DSM 20241<sup>T</sup> indicate that it contains various phospho transferase systems, such as a sucrose-specific EII component (K02808, K02809 and K02810), a β-glucoside β-glucoside-specific EII component (K02755, K02756 and K02757), a cellobiose-specific EII component (K02759, K02760 and K02761), a mannose-specific EII component (K02793, K02794, K02795 and K02796) and an L-ascorbate-specific EII component (K02821, K02822 and K03475) (Fig. 4b), suggesting that strain DSM 20241<sup>T</sup> possesses the ability to ferment various carbon sources.

**Carbon metabolic pathways**
To investigate the fermentative metabolic properties of *L. suionicum* DSM 20241<sup>T</sup>, metabolic pathways of various carbon sources were reconstructed based on predicted KEGG pathways and BLASTP analysis using

---

**Table 3** Sequence features of chromosome and plasmid present in the *L. suionicum* DSM 20241<sup>T</sup> genome

| Label | Size (bp) | Topology | Coding gene sequences (bp) | G + C content (%) | INSDC identifier |
|-------|-----------|----------|---------------------------|------------------|-----------------|
| Chromosome | 2,026,850 | Circular | 1,758,165 | 37.6 | CP015247.1 |
| Plasmid | 21,983 | Circular | 14,895 | 37.0 | CP015248.1 |
reference protein sequences (Fig. 5). The predicted metabolic pathways identified motifs associated with the pentose phosphate pathway, fructose and mannose metabolism, galactose metabolism, sucrose metabolism, pyruvate metabolism, partial TCA cycle and incomplete glycolysis pathway in the genome of *L. suionicum* DSM20241T, indicating that this strain performs typical heterolactic acid fermentation to produce lactate, ethanol and carbon dioxide (Fig. 5, Additional file 1: Table S1). It has been reported that mannitol, an important refreshing sweet agent in fermented vegetable foods such as sauerkraut, pickles and kimchi, is synthesized through fructose reduction by mannitol dehydrogenase (EC 1.1.1.67) through the consumption of NADH [20, 21]. The predicted metabolic pathways indicate that *L. suionicum* DSM20241T produces ethanol via the reduction of acetyl phosphate through the consumption of NADH; this strain may also produce acetate instead of ethanol due to the lack of NADH when the strain produces mannitol from fructose [21]. *L. suionicum* DSM20241T harbors genes related to diverse PTSs or permeases that transport various glycosides or sugars including D-glucose, D-fructose, sucrose, D-mannose, trehalose, arbutin, salcin, cellobiose, D-xylose, arabinose, and D-ribose; this indicates that *L. suionicum* DSM20241T has versatile metabolic capabilities. D-lactate and L-lactate are produced from the reduction of pyruvate by D-lactate dehydrogenase (EC 1.1.1.28) and L-lactate dehydrogenase (EC 1.1.1.27), respectively. *L. suionicum* DSM20241T harbors four copies of D-lactate dehydrogenase (locus tags: Ga0151201_111849, Ga0151201_112070, Ga0151201_11385 and Ga0151201_111758) and one copy of L-lactate dehydrogenase (locus tag: Ga0151201_1175), suggesting that *L. suionicum* DSM20241T may produce more D-lactate than L-lactate; this is similar to other members of the genus *Leuconostoc*, which have been shown to produce more D-lactate than L-lactate under laboratory conditions [4, 22–25]. The predicted metabolic pathways show that *L. suionicum* DSM20241T produces acetylactate synthase (EC 2.2.1.6) produces 2-acetolactate from pyruvate and converts it into deacetyl and CO₂, which is emitted as a byproduct. Furthermore, 2-acetoin is produced from 2-acetolactate and diacetyl (acetolactate decarboxylase, EC 4.1.1.5; diacetyl reductase, EC 1.1.1.304, respectively); but 2-acetoin is eventually converted to 2,3-butanediol, which lacks the butter flavoring property. In addition, the predicted metabolic pathways show that *L. suionicum* DSM20241T uses dextranucrase (EC 2.4.1.5) to produce dextran, a homopolysaccharide of glucose.

**Table 4 Genome statistics**

| Attribute                | Value          | % of Total |
|--------------------------|----------------|------------|
| Genome size (bp)         | 2,048,833      | 100.00     |
| DNA coding (bp)          | 1,835,796      | 89.60      |
| DNA G + C (bp)           | 769,980        | 37.58      |
| DNA scaffolds             | 2              | 100.00     |
| Total genes              | 2090           | 100.00     |
| Protein coding genes     | 1997           | 95.55      |
| RNA genes                | 93             | 4.45       |
| Pseudo genes             | 0              | –          |
| Genes in internal clusters | 381          | 18.23      |
| Genes with function prediction | 1641        | 78.52      |
| Genes assigned to COGs   | 1483           | 70.96      |
| Genes with Pfam domains  | 1695           | 81.10      |
| Genes with signal peptides | 31           | 1.48       |
| Genes with transmembrane helices | 592  | 28.33      |
| CRISPR repeats           | 0              | –          |

**Table 5 Number of genes associated with general COG functional categories**

| Code | Value | %age | Description                                      |
|------|-------|------|--------------------------------------------------|
| J    | 176   | 7.91 | Translation, ribosomal structure and biogenesis  |
| A    | 0     | 0.00 | RNA processing and modification                  |
| K    | 116   | 5.21 | Signal transduction mechanisms                   |
| L    | 83    | 3.73 | Replication, recombination and repair            |
| B    | 0     | 0.00 | Chromatin structure and dynamics                 |
| D    | 25    | 1.12 | Cell cycle control, Cell division, chromosome partitioning |
| V    | 34    | 1.53 | Defense mechanisms                               |
| T    | 53    | 2.38 | Signal transduction mechanisms                   |
| M    | 93    | 4.18 | Cell wall/membrane biogenesis                    |
| N    | 11    | 0.49 | Cell motility                                   |
| U    | 14    | 0.63 | Intracellular trafficking and secretion          |
| O    | 55    | 2.47 | Posttranslational modification, protein turnover, chaperones |
| C    | 53    | 2.38 | Energy production and conversion                 |
| G    | 146   | 6.56 | Carbohydrate transport and metabolism            |
| E    | 179   | 8.04 | Amino acid transport and metabolism              |
| F    | 85    | 3.82 | Nucleotide transport and metabolism              |
| H    | 98    | 4.40 | Coenzyme transport and metabolism                |
| I    | 65    | 2.92 | Lipid transport and metabolism                   |
| P    | 81    | 3.64 | Inorganic ion transport and metabolism           |
| Q    | 27    | 1.21 | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 127   | 5.71 | General function prediction only                 |
| S    | 98    | 4.40 | Function unknown                                 |
| -    | 607   | 27.27| Not in COGs                                     |

The total is based on the total number of protein coding genes in the genome.
Fig. 4 KEGG metabolic (a) and regulatory (b) pathways of *Leuconostoc suionicum* DSM 20241^T. The pathways were generated using the iPath v2 module based on KEGG Orthology numbers of genes identified from the genome of *L. suionicum* DSM 20241^T.
Conclusions

In this study, the complete genome of *L. suionicum* DSM 20241^T^, consisting of a circular chromosome and a circular plasmid, was obtained by whole-genome sequencing using the PacBio SMRT sequencing system and de novo assembly using the HGAP method. In addition, the metabolic pathways of organic compounds in *L. suionicum* DSM 20241^T^ were reconstructed to estimate its fermentative properties and metabolites. The metabolic pathways show that strain DSM 20241^T^ performs typical heterolactic acid fermentations to produce lactate, ethanol and carbon dioxide and contains genes encoding various PTSs, permeases, and other enzymes to metabolize various organic compounds. In addition, strain DSM 20241^T^ synthesizes mannitol to produce acetate instead of ethanol through heterolactic acid fermentation, and produces butter flavoring compounds. The complete genome and reconstructed metabolic pathways of *L. suionicum* DSM 20241^T^ provide important insights into its functional and metabolic features during fermentation.

Additional file

Additional file 1: Table S1. List of genes of *L. suionicum* DSM 20241^T^ in KEGG metabolic pathways. (PDF 411 kb)

Abbreviations

COG: Clusters of orthologous groups; KEGG: Kyoto Encyclopedia of Genes and Genomes; PTS: Phosphotransferase system; SMRT: Single molecule real-time; TCA: Tricarboxylic acid
Funding
This work was supported by the “Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ010900604)” of Rural Development Administration and the World Institute of Kimchi funded by the Ministry of Science, ICT and Future Planning (KE1702–2), Republic of Korea.

Authors’ contributions
BHC and HHJ assembled the sequencing data and completed the genome analysis; SHL and HHJ performed the microbiological studies and obtained the organism information; BHC, SHL and DWK analyzed the genome sequence computationally; BHC and COJ designed the study and wrote the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Department of Life Science, Chung-Ang University, 84, Heuksok-Ro, Dongjak-Gu, Seoul 06974, Republic of Korea. 2Microbiology and Functionality Research Group, World Institute of Kimchi, Gwangju 61755, Republic of Korea. 3Animal Nutrition and Physiology Team, National Institute of Animal Science, RDA, Jeollabukdo 55365, Republic of Korea.

References
1. O’Sullivan L, Ross R, Hill C. Potential of bacteriocin-producing lactic acid bacteria for improvements in food safety and quality. Biochimie. 2002;84:599–604.
2. Pogacíč T, Maillard M-B, Leclerc A, Hervé C, Chuat V, Valence F, Thierry A. Lactobacillus and Leuconostoc volatile compounds in cheese conditions. Appl Microbiol Biotechnol. 2016;100:2335–46.
3. De Bruyne K, Schillinger U, Caroline L, Boehringer B, Cleenwerck I, Vancanneyt M, De Vuyst L, Franz CMAP, Vandamme P. Leuconostoc chung-angensis sp. nov., isolated from sous-vide-cooked rutabaga. Int J Evol Microbiol. 2011;7:2198–71.
4. Hemme D, Foucaud-Scheunemann C. Lactobacillus plantarum subsp. plantarum, isolated from kimchi. Int J Evol Microbiol. 2002;53:1123–9.
5. Lyhs U, Snaauwaert J, Pihalajavita L, De Vuyst L, Vandamme P. Leuconostoc lactis subsp. rapi sp. nov., isolated from sous-vide-cooked cuttage. Int J Evol Microbiol. 2015;65:2586–90.
6. Kim B, Lee J, Jung J, Kim J, Kim H. Leuconostoc inho sp. nov., a lactic acid bacterium isolated from kimchi. Int J Evol Microbiol. 2003;53:1123–6.
7. Cheng H. Volatile flavor compounds in yogurt: a review. Crit Rev Food Sci Nutr. 2010;50:938–50.
8. Hugenholtz J, Keerbeezen M, Starrenburg M, Delcour J, de Vos W, Hols P. Lactococcus lactis as a cell factory for high-level diacetyl production. Appl Environ Microbiol. 2000;66:4112–4.
9. Wang Q, Gartey GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of RNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007;73:5261–7.
10. Felsenstein J. PHYLIP—Phylogenetic inference programs. Ver. 3.68. Computer software and manual. Seattle: University of Washington and Berkeley, University Herbarium, University of California; 1995.
11. Grant Jr, Stothard P. The CoGend server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 2008;36:W181–4.
12. Woese CR, Kandler O, Woese CR. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci. 1990;87:4576–9.
13. Gibbons N, Murray R. Proposals concerning the higher taxa of bacteria. Int J Evol Microbiol. 1978;2:81–8.
14. Garrity GM, Holt JG. The road map to the manual. In: Bergy’s manual of systematic bacteriology, 2nd ed. New York: Springer; 2001. p. 119–66.
15. Ludwig WS, Wirth WH. Class I. The bacilli. In: De Vos P, Gartey GM, Jones D, Krieg NR, Ludwig W, Rainey EA, Schleifer KH, Wirth WH, editors. Bergey’s manual of systematic bacteriology. 2nd ed. New York: Springer; 1990;87:19–20.
16. Ludwig W, Schleifer K, Wirth WH. Order II. Lactobacillales ord. nov. Bergey’s manual of systematic bacteriology, vol. 3. 2nd ed. New York: Springer; 2005. p. 464.
17. Van Tieghem P. Sur la gomme du sucre (Leuconostoc mesenteroides). Ann Sci Nat Bot. 1878;2:180–203.
18. Garvie E. Genus Leuconostoc: van Tieghem 1878, 198AL emend mut. char. hucker and pedersen 1930, 66AL. In: Bergy’s manual of systematic bacteriology. Baltimore: The Williams and Wilkins Co; 1986. p. 1071–5.
19. Skerman VBD, McGowan V, Sneath PHA. Approved lists of bacterial names. Int J Evol Microbiol. 1980;30:225–40.

Page 8 of 8