Genetic diversity of *Lactobacillus delbrueckii* isolated from raw milk in Hokkaido, Japan

H. Tsuchihashi,^{1*} A. Ichikawa,¹ M. Takeda,¹ A. Koizumi,¹ C. Mizoguchi,² T. Ishida,¹ and K. Kimura¹

¹Basic Microbiology Research Department, Food Microbiology Research Laboratories, R&D Division, Meiji Co. Ltd., Hachioji, Tokyo 192-0919, Japan

²Applied Microbiology Research Department, Food Microbiology Research Laboratories, R&D Division, Meiji Co. Ltd., Hachioji, Tokyo 192-0919, Japan

ABSTRACT

Lactic acid bacteria (LAB) play important roles in acid production and flavor formation in fermented dairy products. Lactic acid bacteria strains with distinct characteristics confer unique features to products. Diverse LAB have been identified in raw milk and traditional fermented milk prepared from raw milk. However, little is known about LAB in raw milk in Japan. To preserve diverse LAB as potential starters or probiotics for future use, we have isolated and identified various kinds of LAB from raw milk produced in Japan. In this study, we focused on *Lactobacillus delbrueckii*, one of the most important species in the dairy industry. We identified *L. delbrueckii* subspecies isolated from raw milk in Hokkaido, Japan, by analyzing intraspecific diversity using 4 distinct methods, *hsp60* cluster analysis, multilocus sequence analysis, core-genome analysis, and whole-genome analysis based on average nucleotide identity. The subspecies distribution and a new dominant subset of *L. delbrueckii* from raw milk in Japan were revealed. The discovery of new strains with different genotypes is important for understanding the geographic distribution and characteristics of the bacteria and further their use as a microbial resource with the potential to express unconventional flavors and functionalities. The strains identified in this study may have practical applications in the development of fermented dairy products.

Key words: *Lactobacillus delbrueckii*, *hsp60* (*groEL*), multilocus sequence analysis (MLSA), core-genome analysis, average nucleotide identity (ANI)

INTRODUCTION

Lactic acid bacteria (LAB) are used in the production of various fermented foods, especially fermented

dairy products. To meet the diverse needs of consumers, LAB strains with various characteristics have been developed as starters and are used in the production of dairy products worldwide. Traditional dairy products such as artisan cheese contain various LAB originating from not only natural starter cultures but also raw milk used as a base material (Callon et al., 2004; Poznanski et al., 2004; Coppola et al., 2006). In contrast, industrially fermented dairy products are manufactured using a starter consisting of a small number of strains and pasteurized milk. Although this is the most effective strategy to control quality and ensure safety, it tends to result in poor or weak flavors compared with traditional fermented dairy products (Wijesundera et al., 1997; Fox et al., 1998). This mainly reflects the low diversity of LAB in pasteurized milk and starters, which contain limited species or strains (Wouters et al., 2002; Van Hoorde et al., 2010a). There are several reports on the composition and diversity of LAB derived from raw milk (Giannino et al., 2009; Van Hoorde et al., 2010b). Japan has various traditional fermented foods, but no traditional fermented dairy products; hence, data on indigenous LAB in raw milk in Japan are limited (Hagi et al., 2010, 2013).

The intraspecific diversity of LAB isolated from raw milk has been evaluated using random amplified polymorphic DNA (RAPD; Moschetti et al., 1998; Mora et al., 2002; Giraffa et al., 2004). Random amplified polymorphic DNA is commonly used for bacterial typing owing to its convenience; however, it is limited by low reproducibility. Recently, analyses of housekeeping genes have been used as a complementary method to RAPD. The *hsp60* (*groEL*) gene is useful for identifying bacterial subspecies that cannot be identified using 16S rDNA and for elucidating diversity within species (Zeaiter et al., 2002; Masco et al., 2004; Avendaño-Herrera et al., 2014). Multilocus sequence typing or multilocus sequence analysis (MLSA), based on 7 or 8 housekeeping genes, is an alternative approach to identify subspecies (Tanigawa and Watanabe, 2011) and characterize intraspecific diversity (Yu et al., 2015;

Received August 9, 2021.

Accepted November 3, 2021.

*Corresponding author: hanae.tsuchihashi@meiji.com

Delorme et al., 2017), intrasubspecific diversity (Liu et al., 2016; Song et al., 2016), and regional diversity of intersubspecies (Song et al., 2016). Moreover, next-generation sequencing techniques have been used to sequence whole bacterial genomes for comparative analyses of closely related strains (Fontana et al., 2018; Frantzen et al., 2018; Alexandraki et al., 2019). Average nucleotide identity (ANI) is an index developed to distinguish species instead of DNA-DNA hybridization, which is complicated to operate. Average nucleotide identity is widely used in species identification as a standard approach (Chun et al., 2018). Although an ANI of 95 to 96% corresponds to DNA-DNA hybridization 70% (Konstantinidis and Tiedje, 2005; Goris et al., 2007; Richter and Rosselló-Móra, 2009), there is no ANI threshold to distinguish subspecies. However, recently, there have been reports on the use of ANI to assess diversity within species (Tanizawa et al., 2016) and to distinguish subspecies (Li et al., 2021).

Lactobacillus delbrueckii is one of the most important species in the dairy industry; it has 6 subspecies (Weiss et al., 1983; Dellaglio et al., 2005; Kudo et al., 2012; Adimpong et al., 2013), including *Lactobacillus delbrueckii* ssp. *bulgaricus*, a yogurt starter. As the subspecies are indistinguishable by 16S rDNA sequences (Stackebrandt, 2006), subspeciation of *L. delbrueckii* has been performed by multilocus sequence typing (Tanigawa and Watanabe, 2011). Despite the frequent isolation of this species from various fermented foods and analyses of diversity (Giraffa et al., 2004; Mohammed et al., 2009; Song et al., 2016), only a few studies have focused on intraspecific genetic diversity based on isolates from raw milk.

Recently, various fermented dairy products have been developed at both local small-scale and industrial scale in Japan. To expand the dairy industry in Japan, it is necessary to preserve diverse LAB in raw milk as potential starters or probiotics for future production of unique dairy products. Therefore, we have isolated and identified various kinds of LAB, including *L. delbrueckii*, from raw milk in Japan since 2009.

The objective of this study is to identify the subspecies of 226 *L. delbrueckii* strains isolated from raw milk in Hokkaido, the largest dairy farming region in Japan, from 2009 to 2015, and estimate their distribution. For species classification, polyphasic analysis with multiple indicators is an effective and reliable method, since the monophasic analysis always has potential errors (Vandamme et al., 1996). In this study, the intraspecific diversity of *L. delbrueckii* derived from raw milk was evaluated using 4 distinct methods, *hsp60* cluster analysis, MLSA, core-genome analysis, and whole-genome analysis based on ANI.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

A total of 228 *L. delbrueckii* strains including 226 strains derived from raw milk produced in Hokkaido were used in this study. The 226 strains were obtained through 6 different studies on the isolation of LAB from raw milk conducted from 2009 to 2015. The raw milk from which these strains were isolated was sampled from bulk tanks at several plants that belong to Meiji Company Ltd. (formerly Meiji Dairies Corporation) in Hokkaido. The strains were isolated using multiple types of media preferred by LAB such as de Man, Rogosa, and Sharpe, M17, and Lactobacillus Selection (LBS) agar and identified as *L. delbrueckii* by partial 16S ribosomal DNA sequencing, as described previously (Kimura et al., 2010). These 226 strains were confirmed to be free of duplicates by RAPD. *Lactobacillus delbrueckii* ssp. *sunkii* NRIC 0766 and NRIC 0767 were obtained from the Center for Microorganism Resources at the Tokyo University of Agriculture. All strains were grown in de Man, Rogosa, and Sharpe broth (BD Difco) at 37°C for 18 h.

Sequencing of 7 Housekeeping Genes

Bacterial DNA was extracted from each isolate using the InstaGene Matrix (Bio-Rad Laboratories Inc.) according to the manufacturer's instructions. Seven housekeeping genes (*fusA*, *gyrB*, *hsp60*, *ileS*, *pyrG*, *recA*, and *recG*) were sequenced using the primers and the amplification program reported by Tanigawa and Watanabe (2011). The reaction mixture (20 µL) was composed of 0.2 µL of TaKaRa Ex Taq (Takara Bio), 2 µL of 10 × Ex Taq Buffer (Takara Bio), 1.6 µL of dNTP mixture (2.5 mM each, Takara Bio), 0.6 µL of each primer (5 pmol/µL), 0.8 µL of template DNA, and 14.2 µL of distilled water. The PCR products were purified and sequenced as described by Kimura et al. (2010).

First, *hsp60* (511 bp) of the 226 *L. delbrueckii* strains isolated from raw milk from the Hokkaido region was sequenced for a cluster analysis. Thereafter, 26 representative strains including all novel alleles were selected. Seven housekeeping genes were sequenced in 30 strains, and the NODAI Culture Collection (NRIC) and ME numbers are shown in Table 1.

Cluster Analysis of Hsp60

A cluster analysis of the *hsp60* alleles found in the 226 strains was performed using the neighbor-joining

Table 1. *Lactobacillus delbrueckii* strains, genome sequences, and cluster analysis results

Species, subspecies, strain no. (origin)	Genome accession no.	<i>hsp60</i> allele no.	<i>hsp60</i> group	MLSA ¹ cluster	Reference
<i>L. delbrueckii</i> ssp. <i>lactis</i>					
DSM 20072 ^T	CP022988.1	3	1	I-B2	Weiss et al., 1983
KCCM 34717	CP018215.1	3	1	I-B2	—
KCTC 3034	CP023139.1	3	1	I-B2	—
KCTC 3035	NZ_CP018156.1	4	1	I-A	—
Lactis1	NZ_LS991409.1	15	1	I-D	—
NWC_1_2	CP029250.1	3	1	I-B2	—
MAG_rmk202	CP046131.1	3	1	I-B2	—
<i>Lactobacillus delbrueckii</i> ssp. <i>jakobsenii</i>					
DSM 26046 ^T	NZ_CP018218.1	3	1	I-C	Adimpong et al., 2013
<i>Lactobacillus delbrueckii</i> ssp. <i>delbrueckii</i>					
DSM 20074 ^T = JCM 1012 ^T	NZ_CP018615.1	1	1	II	Weiss et al., 1983
KCTC 13731	NZ_CP018216.1	28	1	I-G	—
TUA4408L	NZ_CP021136.1	3	1	I-C	—
<i>L. delbrueckii</i> ssp. <i>sunkii</i>					
JCM 17838 ^T	NZ_CP018217.1	11	3	III-A	Kudo et al., 2012
NRIC 0766 (Japanese traditional pickle “sunki”)	BNHN	11	3	III-A	This study
NRIC 0767 (Japanese traditional pickle “sunki”)	BNHO	11	3	III-A	This study
<i>L. delbrueckii</i> ssp. <i>bulgaricus</i>					
ATCC 11842 ^T = JCM 1002 ^T	NC_008054.1	5	2	IV	Weiss et al., 1983
2038	NC_0174669.1	5	2	IV	—
<i>L. delbrueckii</i> ssp. <i>indicus</i>					
JCM 115610 ^T	NZ_CP018614.1	9	4	V	Dellaglio et al., 2005
ME-781 (yogurt, Nepal)	—	23	4	V	This study
<i>L. delbrueckii</i>					
ME-782 (raw milk, Hokkaido, Japan)	BNHP	11	3	III-B	This study
ME-783 (raw milk, Hokkaido, Japan)	BNHQ	11	3	III-B	This study
ME-784 (raw milk, Hokkaido, Japan)	BNHR	15	1	I-B2	This study
ME-785 (raw milk, Hokkaido, Japan)	BNHS	15	1	I-B2	This study
ME-786 (raw milk, Hokkaido, Japan)	BNHT	18	1	I-B1	This study
ME-787 (raw milk, Hokkaido, Japan)	BNHU	18	1	I-B1	This study
ME-788 (raw milk, Hokkaido, Japan)	BNHV	26	1	I-B1	This study
ME-789 (raw milk, Hokkaido, Japan)	BNHW	19	1	I-D	This study
ME-790 (raw milk, Hokkaido, Japan)	BNHX	4	1	I-A	This study
ME-791 (raw milk, Hokkaido, Japan)	BNHY	11	3	III-B	This study
ME-792 (raw milk, Hokkaido, Japan)	BNHZ	21	4	V	This study
ME-793 (raw milk, Hokkaido, Japan)	BNIA	4	1	I-B1	This study
ME-794 (raw milk, Aomori, Japan)	—	25	3	III-B	This study
ME-795 (raw milk, Hokkaido, Japan)	BNIB	4	1	I-B1	This study
ME-796 (raw milk, Hokkaido, Japan)	BNIC	24	3	III-B	This study
ME-797 (raw milk, Hokkaido, Japan)	BNID	20	1	I-E	This study
ME-798 (raw milk, Hokkaido, Japan)	BNIE	4	1	I-B1	This study
ME-799 (raw milk, Hokkaido, Japan)	BNIF	18	1	I-B1	This study
ME-800 (raw milk, Hokkaido, Japan)	BNIG	11	3	III-B	This study
ME-801 (raw milk, Hokkaido, Japan)	BNIH	4	1	I-B2	This study
ME-802 (raw milk, Hokkaido, Japan)	BNII	14	1	I-C	This study
ME-803 (raw milk, Hokkaido, Japan)	BNIJ	18	1	I-B1	This study
ME-804 (raw milk, Hokkaido, Japan)	BNIK	27	1	I-B1	This study
ME-805 (raw milk, Hokkaido, Japan)	BNIL	18	1	I-B1	This study
ME-806 (raw milk, Hokkaido, Japan)	BNIM	18	1	I-B1	This study
ME-807 (raw milk, Hokkaido, Japan)	BNIN	22	1	I-B2	This study
ME-808 (raw milk, Hokkaido, Japan)	BNIO	15	1	I-D	This study

¹MLSA = multilocus sequence analysis.

method (Saitou and Nei, 1987) with Genetyx Ver. 13 (Genetyx Co.). Seventeen *hsp60* sequences (hsp60-1 to hsp60-17) previously reported by Tanigawa and Watanabe (2011) and 3 *hsp60* sequences (hsp60-23, hsp60-25, and hsp60-28) of strains with different origins (ME-781, ME-794, and KCTC 13731; Table 1) were included in the cluster analysis.

MLSA

All 7 housekeeping genes were concatenated into single sequences. A total of 86 concatenated sequences, consisting of 45 sequences obtained from strains shown in Table 1 and 41 previously reported sequences (Tanigawa and Watanabe, 2011), were used for MLSA.

Sequence alignment and phylogenetic analyses of 80 concatenated sequences (length: 3,155 bp) were performed using MEGA X version 10.0.5 (Kumar et al., 2018). Evolutionary distances were computed using the Tamura 3-parameter method (Tamura, 1992). A phylogenetic tree with 1,000 bootstrap replicates (Felsenstein, 1985) was constructed using the unweighted pair group method with arithmetic mean (UPGMA) method (Sneath and Sokal, 1973).

Extraction and Purification of Genomic DNA from *L. delbrueckii*

For the 28 strains with NRIC and ME numbers, excluding ME-781 and ME-794, listed in Table 1, genomic DNA was extracted for genome sequencing.

One milliliter of each culture was used for genomic DNA extraction. Extraction and purification of genomic DNA were performed using 2 column-based commercial kits, the Wizard Genomic DNA Purification Kit (A1220; Promega) and Genomic DNA Clean and Concentrator Kit (#D4011, Zymo Research). Before purification, the cells were collected by centrifugation [at $13,000 \times g$ for 2 min at room temperature (24°C)], washed once with an equal volume of 5 mM Tris-HCl (pH 8.0), and suspended in 100 µL of 20 mM Tris-HCl (pH 7.0) buffer containing 0.3 mM raffinose, 5 mM CaCl₂, and 5 mM MgCl₂. Cell wall digestion was performed by adding 20 µL of 10 mg/mL lysozyme (L6876, chicken egg white; Sigma-Aldrich) and 5 µL of 10 U/µL mutanolysin (catalog number 1017-10; A&A Biotechnology) and incubating at 37°C for 10 min. Pellets collected by centrifugation after digestion were suspended in 600 µL of Nuclei Lysis solution provided in the Wizard Genomic DNA Purification Kit and mixed gently. The first purification was performed according to the manufacturer's instructions. Thereafter, 100 µL of genomic DNA was suspended in 200 µL of Chip DNA Binding Buffer (Genomic DNA Clean and Concentrator). The second purification was performed according to the manufacturer's instructions. Finally, genomic DNA was eluted with 100 µL of sterilized water and quantified using the Qubit dsDNA HS Assay Kit (Q32851; Thermo Fisher Scientific).

DNA Library Preparation and Genome Sequencing

The DNA library was prepared using the QIAseq FX DNA Library Kit (#180473; Qiagen) according to the manufacturer's instructions. Briefly, 100 ng of DNA was used as the input and fragmented at 32°C for 10 min. After adapter ligation, the sample was purified using Agencourt AMPure XP (Beckman Coulter, Inc.). The quality and quantity of DNA were assessed using the

Agilent BioAnalyzer and QIAseq Library Quant Assay Kit (#333314; Qiagen), respectively. The library (0.5 nM) was denatured and diluted to 10 pM (phiX: 1% spike in) following the NextSeq System Denature and Dilute Libraries Guide (Illumina Inc.). Genomes were sequenced using the Illumina MiSeq platform.

Genome Assembly, Prediction, and Annotation of Coding Sequences

Twenty-eight draft genomes were assembled using Velvet (Zerbino and Birney, 2008) with the VelvetOptimizer script (<https://github.com/tseemann/VelvetOptimiser>). The coding sequences of 43 genome sequences shown in Table 1 (except for ME-781 and ME-794) were predicted and annotated using DFAST v. 1.1.0 (Tanizawa et al., 2016, 2018). The protein.faa files generated from DFAST were used as the input files for the core-genome analysis using BPGA v. 1.3 (Chaudhari et al., 2016).

Core-Genome Analysis

The amino acid sequences of 529 core genes extracted with an identity cut-off of 0.97 using USE-ARCH v. 11.0.667 (Edgar, 2010) built in BPGA were concatenated using CONCAT.sh (<https://github.com/abhijeetsingh1704/CONCAT/>) and aligned using MAFFT (Katoh et al., 2002) v. 7.215 (Katoh and Standley, 2013) with the *-auto* option. A phylogenetic analysis of 43 concatenated AA sequences (average length: $167,246 \pm 635$ AA) was performed using MEGA X v. 10.0.5 (Kumar et al., 2018). Evolutionary distances were computed using the Poisson correction method (Zuckerandl and Pauling, 1965). A phylogenetic tree with 1,000 bootstrap replicates (Felsenstein, 1985) was constructed using the UPGMA method (Sneath and Sokal, 1973).

Whole-Genome Analysis Based on ANI

OrthoANIu tool (Yoon et al., 2017) was used to calculate ANI; the distance (1-ANI/100) was calculated from ANI (%) to construct the dendrogram, and R 4.0.2 (<https://www.r-project.org/>) was used to construct the dendrogram using the UPGMA method (Sneath and Sokal, 1973).

Data Availability

The DDBJ/ENA/GenBank accession numbers for the 7 housekeeping gene sequences of 30 strains with NRIC and ME numbers listed in Table 1 are LC578537–LC578566 (*fusA*), LC578567–LC578596

(*gyrB*), LC578597–LC578626 (*hsp60*), LC578627–LC578656 (*ileS*), LC578657–LC578686 (*pyrG*), LC578687–LC578716 (*recA*), and LC578717–LC578746 (*recG*). The genome sequences obtained in this study are deposited under the BioProject PRJDB10495.

RESULTS AND DISCUSSION

Diversity of *Hsp60* Alleles of the 226 *L. delbrueckii* Strains

The *hsp60* of the 226 *L. delbrueckii* strains isolated from raw milk in Hokkaido was sequenced. Six previously reported alleles (3, 4, 7, 11, 14, and 15) and 8 novel alleles (no. 18–22, 24, 26, and 27) were identified (Table 2). Of the 226 strains tested, 126 (55.8%) had allele number 18.

The nucleotide identity percentage among 28 *hsp60* sequences is shown in Supplemental Table S1 (<https://doi.org/10.5281/zenodo.5650124>). All 28 *hsp60* sequences were clustered into 4 groups according to the subspecies (Figure 1). Group 1 included 2 alleles: number 1 was carried by the type strain of *L. delbrueckii* ssp. *delbrueckii* and number 3 was shared by 2 type strains of *L. delbrueckii* ssp. *lactis* and *L. delbrueckii* ssp. *jakobsenii*. Group 2 consisted of only alleles car-

ried by *L. delbrueckii* ssp. *bulgaricus*. Groups 3 and 4 included alleles of the type strain of *L. delbrueckii* ssp. *sunkii* and *L. delbrueckii* ssp. *indicus*, respectively.

Alleles belonging to groups 1 and 3 were carried by 199 (88.1%) and 26 strains (11.5%), respectively. Only one strain carrying allele number 21 belonged to group 4. No strains carried the allele belonging to group 2 among the 226 strains tested.

These results suggest that strains carrying the *hsp60* allele belonging to group 1, especially allele number 18, were dominant among *L. delbrueckii* strains derived from raw milk in Japan. We also examined the *hsp60* allele in nearly 200 isolates from raw milk procured outside Hokkaido and confirmed that the distribution of the *hsp60* allele is not limited to raw milk from Hokkaido (data not shown). We are interested in determining whether this distribution of *hsp60* alleles and the existence of allele number 18 are unique to the *L. delbrueckii* strain derived from raw milk in Japan.

MLSA of *L. delbrueckii* Isolated from Raw Milk in Japan

To estimate the distribution of subspecies among the 226 strains isolated from raw milk in Japan, 26 representative strains including all novel alleles were selected for subspecies identification by MLSA (Tanigawa and Watanabe, 2011). All clusters or subclusters generated by MLSA were newly named because they were reorganized in the context of newly isolated strains.

Five main clusters (I to V, Figure 2) corresponding to 5 subspecies (*lactis*, *delbrueckii*, *sunkii*, *bulgaricus*, and *indicus*) were identified in the phylogenetic analysis of 7 concatenated housekeeping genes. *Lactobacillus delbrueckii* ssp. *jakobsenii* belonged to cluster I. These results were consistent with those of previous studies (Tanigawa and Watanabe, 2011; Kudo et al., 2012; Adimpong et al., 2013) and suggested that strains assigned to cluster I cannot be identified as *lactis* or *jakobsenii* species. Therefore, we used cluster, subcluster, or subset names in this study.

Of the 26 strains derived from raw milk, 25 were classified into cluster I or III. Cluster I was divided into 7 subclusters (from I-A to I-G), and subcluster I-B consisted of 2 subsets (I-B1 and I-B2). The subclusters I-B2 and I-C included type strains of *L. delbrueckii* ssp. *lactis* and *L. delbrueckii* ssp. *jakobsenii*, respectively. Strains derived from raw milk were classified into 5 subclusters (from I-A to I-E), and subcluster I-B1 consisted of only strains derived from raw milk. Interestingly, *hsp60* allele number 18, the most dominant allele, was found only in subcluster I-B1.

Cluster III was divided into 2 subclusters (III-A and III-B). Subcluster III-A consisted only *L. delbrueckii*

Table 2. Distribution of the *hsp60* alleles, *hsp60* group, and multilocus sequence analysis (MLSA) cluster among 226 strains of *Lactobacillus delbrueckii* isolated from raw milk in Hokkaido, Japan

<i>hsp60</i> allele number	Number of strains tested	Proportion (%)	<i>hsp60</i> group	MLSA cluster
18	126	55.8	1	I
4	50	22.1	1	I
15	9	4.0	1	I
19	5	2.2	1	I
13 or 14	4	1.8	1	I
3	1	0.4	1	I
20	1	0.4	1	I
22	1	0.4	1	I
26	1	0.4	1	I
27	1	0.4	1	I
1	0	0.0	1	II
2	0	0.0	1	I
10	0	0.0	1	I
16	0	0.0	1	I
17	0	0.0	1	I
28	0	0.0	1	I
5	0	0.0	2	IV
6	0	0.0	2	IV
8	0	0.0	2	IV
11	22	9.7	3	III
24	3	1.3	3	III
7	1	0.4	3	III
25	0	0.0	3	III
21	1	0.4	4	V
9	0	0.0	4	V
12	0	0.0	4	V
23	0	0.0	4	V
Total	226	—	—	—

ssp. *sunkii*, whereas III-B consisted of strains isolated from raw milk and YIT 0373. Strains belonging to III-B were not named *L. delbrueckii* ssp. *sunkii* in this study because these strains could utilize lactose except for ME-800 (data not shown) contrary to the description of *L. delbrueckii* ssp. *sunkii*, and YIT 0373 was not included in *L. delbrueckii* ssp. *sunkii* in a previous study (Kudo et al., 2012). The relationship between the lactose utilization ability and subclusters found in III-A and III-B is consistent with previous findings (Tanigawa and Watanabe, 2011). Strains belonging to III-A and

ME-800 (III-B) showed very low homology against both lactose permease gene (*lacS*) and β -galactosidase gene (*lacZ*) of *L. delbrueckii* ssp. *bulgaricus* ATCC 11842^T, suggesting mutation or deletion of the lactose operon (data not shown). As the sugar utilization ability reflects the origin of isolates, differences in subclusters within the same cluster may therefore reflect the origin rather than differences in subspecies.

Only one strain (ME-792) derived from raw milk was classified into cluster V and was named *L. delbrueckii* ssp. *indicus*. To the best of our knowledge, this is the

Figure 1. Dendrogram based on *hsp60* sequence. The scale bar represents 0.001 substitutions per site. The numbers in parentheses indicate the number of strains detected in 226 strains of *Lactobacillus delbrueckii* derived from raw milk.

Figure 2. Phylogenetic tree based on the concatenated nucleotide sequences of 7 housekeeping genes. The optimal tree (sum of branch lengths = 0.08713860) is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates, threshold: 70%) are shown next to the branches. The scale bar represents 0.001 substitutions per site.

first report of *L. delbrueckii* ssp. *indicus* isolated from raw milk in Japan. The discovery of this strain, belonging to *L. delbrueckii* ssp. *indicus*, previously reported in fermented milk in India (Dellaglio et al., 2005) and also found by us in fermented milk in Nepal (ME-781), in Japanese raw milk may have important implications for the geographical comparison of distribution and genetic diversity.

The *hsp60* cluster analysis and MLSA showed that MLSA has a higher resolution than cluster analysis using a single *hsp60* gene in our study, as described by Tanigawa and Watanabe (2011). Although a single gene analysis using *hsp60* is inferior to MLSA in terms of its ability to identify strains, the results suggest that *hsp60* cluster analysis may be useful as a simple method to estimate the subspecies of *L. delbrueckii*.

Core-Genome Analysis of *L. delbrueckii*

The *hsp60* cluster analysis and MLSA showed that subcluster I-B1 is the dominant subset in *L. delbrueckii* strains derived from raw milk in Hokkaido. However, as evidenced by the lower bootstrap values, subcluster I-B1 and cluster I were not well supported compared with other clusters, such as IV and V (Figure 2). These results suggest that the 7 genes were not sufficient to clearly distinguish *L. delbrueckii* belonging to cluster I. We performed a core-genome analysis using the BPGA tool with 43 strains, including 17 strains as a reference.

The topology of the tree based on 529 core genes (Figure 3) was very similar to that based on 7 housekeeping genes (Figure 2); the bootstrap values of the subclusters within cluster I improved substantially. The core-genome analysis demonstrated that there were subsets within cluster I, and strains belonging to sub-cluster I-B1 consistently formed an independent sub-cluster supported by a high bootstrap value (99.5%).

Whole-Genome Analysis Based on ANI

The ANI is widely used in species identification as a standard approach, although not in subspecies identification (Chun et al., 2018). However, Tanizawa et al. (2016) suggested, using the genome data of *L. delbrueckii*, that ANI is reliable for the evaluation of intra-specific diversity; therefore, we evaluated our strains using ANI.

The ANI ranged from 97.00% to 99.96%, suggesting that the strains were closely related. The dendrogram based on the ANI (Figure 4) also presented a topology similar to that based on 7 housekeeping genes (Figure 2) and core-genome analysis (Figure 3). The reason for the similarity in the results of the core-genome analysis and those of the ANI-based analysis is that both

Figure 3. Phylogenetic tree based on the concatenated amino acid sequences of 529 core genes. The optimal tree (sum of branch lengths = 0.07864525) is shown. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates, threshold: 70%) are shown next to the branches. The scale bar represents 0.001 substitutions per site.

Figure 4. Dendrogram based on the average nucleotide identity (ANI). The scale bar represents $1 - \text{ANI} (\%)/100$ as distance. The distance from 0.000 to 0.030 means from ANI 100% to 97%, respectively.

analyses were conducted on orthologous genes. The consistency of the results of genome-based analysis and MLSA indicates that the 7 housekeeping genes are sufficient for the subspeciation of *L. delbrueckii*.

Of the 4 different analyses, MLSA was the most reliable and simplest to operate in identifying the subspecies of *L. delbrueckii*.

Distribution of MLSA Clusters (Subspecies) Among the 226 Strains

We focused on the relationship between the *hsp60* alleles and MLSA clusters to estimate the distribution of subspecies in 226 strains derived from raw milk. All strains carrying alleles belonging to group 1 (except for allele number 1) based on *hsp60* sequences were classified by MLSA into cluster I, which included the type strains of *L. delbrueckii* ssp. *lactis* and *jakobsenii*. In addition, other groups in the *hsp60* cluster corresponded to the MLSA cluster one-on-one without crossover among different MLSA clusters (Tables 1 and 2). Accordingly, the distribution of the MLSA cluster (nearly equal to subspecies) was estimated as I: 88.1%, II: 0.0%, III: 11.5%, IV: 0.0%, and V: 0.4% (Table 2).

Although there have been reports of *L. delbrueckii* in fermented foods (Giraffa et al., 2004; Mohammed et al., 2009) and *L. delbrueckii* ssp. *bulgaricus* (Song et al., 2016), there have been no detailed analyses of *L.*

delbrueckii in raw milk in other areas. Therefore, the intraspecific diversity observed in this study cannot be compared with previous findings. However, the bacterial flora in raw milk collected in tanks on farms after milking is affected by the season, farm environment, cow hygiene and rearing environment, and storage conditions of raw milk. *Lactobacillus delbrueckii* evaluated in this study was isolated from milk sampled from a tank in a milk plant, and its flora may be influenced by various factors (Hagi et al., 2010; Hagi et al., 2013; Cao et al., 2021; Hornik et al., 2021; Nikoloudaki et al., 2021).

Three of the 5 strains carrying *hsp60* allele number 4 and all 8 strains carrying number 18, 26, or 27 were assigned to subcluster I-B1 by MLSA of 26 strains. It was estimated that among the 226 strains, 199 strains (69.9%) belonged to subcluster I-B1. Although subcluster I-B1 is not an independent subspecies, it was the largest subset, accounting for 69.9% of *L. delbrueckii* strains derived from raw milk. Finally, these results suggest that strains assigned to cluster I, specifically subcluster I-B1, were dominant among *L. delbrueckii* isolated from raw milk in Hokkaido.

The strains belonging to subcluster I-B1 (I-B1 strain) can potentially impart different flavor characteristics compared with the commercial yogurt starter *L. delbrueckii* ssp. *bulgaricus* strain. In the future, the characterization of I-B1 strain will help to develop fermented

dairy products with different flavors from the yogurt starter. However, the ability of the I-B1 strain to ferment milk is slightly weaker than that of yogurt starter (data not shown). Therefore, it would be more effective to use it in combination with a strain having a superior ability to ferment milk, such as *S. thermophilus* strains possessing *prtS*, which encodes the cell wall protease (Shahbal et al., 1991), when producing yogurt industrially (Yamamoto et al., 2020).

CONCLUSIONS

To the best of our knowledge, this study was the first to demonstrate the intraspecific genetic diversity of *L. delbrueckii* derived from raw milk in Hokkaido. We identified the subspecies and estimated their distribution within the 226 strains isolated from raw milk produced in Hokkaido using 4 distinct approaches (*hsp60* cluster analysis, MLSA, core-genome analysis, and whole-genome analysis based on ANI). The distribution of the MLSA cluster (nearly equal to subspecies) was estimated as I: 88.1%, II: 0.0%, III: 11.5%, IV: 0.0%, and V: 0.4%. One strain of *L. delbrueckii* ssp. *indicus* (cluster V) was discovered for the first time in raw milk in Japan. Moreover, we found that subcluster I-B1, which comprised 66.9% of the total strains, was the most dominant subset of *L. delbrueckii* derived from raw milk in Hokkaido. The discovery of new strains with different genotypes is important for understanding the behavior and geographic distribution of the bacteria and using them as a microbial resource with potential to impart unconventional flavors and functionalities. In the future, characterizing the strains derived from raw milk including I-B1 subsets and elucidating their differences from the yogurt starter *L. delbrueckii* ssp. *bulgaricus* (cluster IV) may have practical applications in the development of new dairy products such as yogurt.

ACKNOWLEDGMENTS

The authors are grateful to Mitsuo Sakamoto (RIKEN BioResource Research Center, Tsukuba, Japan) for helpful discussions. The authors also thank Seiya Makino (Meiji Co., Ltd., Hachioji, Japan) for providing constructive comments and offering encouragement. The authors have not stated any conflicts of interest.

REFERENCES

- Adimpong, D. B., D. S. Nielsen, K. I. Sorensen, F. K. Vogensen, H. Sawadogo-Lingani, P. M. Derkx, and L. Jespersen. 2013. *Lactobacillus delbrueckii* ssp. *jakobsenii* ssp. nov., isolated from dolo wort, an alcoholic fermented beverage in Burkina Faso. *Int. J. Syst. Evol. Microbiol.* 63:3720–3726. <https://doi.org/10.1099/ijls.0.048769-0>.
- Alexandraki, V., M. Kazou, J. Blom, B. Pot, K. Papadimitriou, and E. Tsakalidou. 2019. Comparative genomics of *Streptococcus thermophilus* support important traits concerning the evolution, biology and technological properties of the species. *Front. Microbiol.* 10:2916. <https://doi.org/10.3389/fmicb.2019.02916>.
- Avendaño-Herrera, R., S. Balboa, N. Castro, A. Gonzalez-Contreras, B. Magarinos, J. Fernandez, A. E. Toranzo, and J. L. Romalde. 2014. Comparative polyphasic characterization of *Streptococcus phocae* strains with different host origin and description of the subspecies *Streptococcus phocae* ssp. *salmonis* ssp. nov. *Int. J. Syst. Evol. Microbiol.* 64:1775–1781. <https://doi.org/10.1099/ijls.0.056978-0>.
- Callon, C., L. Millet, and M. C. Montel. 2004. Diversity of lactic acid bacteria isolated from AOC Salers cheese. *J. Dairy Res.* 71:231–244. <https://doi.org/10.1017/S0022029904000159>.
- Cao, H., Y. Yan, L. Wang, L. Dong, X. Pang, S. Tang, A. Li, A. Xiang, L. Zhang, and B. Zheng. 2021. High-throughput sequencing reveals bacterial diversity in raw milk production environment and production chain in Tangshan City of China. *Food Sci. Anim. Resour.* 41:452–467. <https://doi.org/10.5851/kosfa.2021.e10>.
- Chaudhari, N. M., V. K. Gupta, and C. Dutta. 2016. BPGA-An ultra-fast pan-genome analysis pipeline. *Sci. Rep.* 6:24373. <https://doi.org/10.1038/srep24373>.
- Chun, J., A. Oren, A. Ventosa, H. Christensen, D. R. Arahall, M. S. da Costa, A. P. Rooney, H. Yi, X. W. Xu, S. De Meyer, and M. E. Trujillo. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68:461–466. <https://doi.org/10.1099/ijsem.0.002516>.
- Coppola, S., V. Fusco, R. Andolfi, M. Aponte, G. Blaiotta, D. Ercolini, and G. Moschetti. 2006. Evaluation of microbial diversity during the manufacture of Fior di Latte di Agerola, a traditional raw milk pasta-filata cheese of the Naples area. *J. Dairy Res.* 73:264–272. <https://doi.org/10.1017/S0022029906001804>.
- Dellaglio, F., G. E. Felis, A. Castioni, S. Torriani, and J. E. Germond. 2005. *Lactobacillus delbrueckii* ssp. *indicus* ssp. nov., isolated from Indian dairy products. *Int. J. Syst. Evol. Microbiol.* 55:401–404. <https://doi.org/10.1099/ijls.0.63067-0>.
- Delorme, C., N. Legravet, E. Jamet, C. Hoarau, B. Alexandre, W. M. El-Sharoud, M. S. Darwish, and P. Renault. 2017. Study of *Streptococcus thermophilus* population on a world-wide and historical collection by a new MLST scheme. *Int. J. Food Microbiol.* 242:70–81. <https://doi.org/10.1016/j.jfoodmicro.2016.11.016>.
- Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>.
- Fontana, A., C. Zaccari, and L. Morelli. 2018. Genetic signatures of dairy *Lactobacillus casei* group. *Front. Microbiol.* 9:2611. <https://doi.org/10.3389/fmicb.2018.02611>.
- Fox, P. F., P. L. H. McSweeney, and C. M. Lynch. 1998. Significance of non-starter lactic acid bacteria in cheddar cheese. *Aust. J. Dairy Technol.* 53:83–89.
- Frantzen, C. A., H. P. Kleppen, and H. Holo. 2018. *Lactococcus lactis* diversity in undefined mixed dairy starter cultures as revealed by comparative genome analyses and targeted amplicon sequencing of *epsD*. *Appl. Environ. Microbiol.* 84. <https://doi.org/10.1128/AEM.02199-17>.
- Giannino, M. L., M. Marzotto, F. Dellaglio, and M. Feligini. 2009. Study of microbial diversity in raw milk and fresh curd used for Fontina cheese production by culture-independent methods. *Int. J. Food Microbiol.* 130:188–195. <https://doi.org/10.1016/j.jfoodmicro.2009.01.022>.
- Giraffa, G., C. Andrighetto, C. Antonello, M. Gatti, C. Lazzi, G. Marcuzzan, A. Lombardi, and E. Neviani. 2004. Genotypic and phenotypic diversity of *Lactobacillus delbrueckii* ssp. *lactis* strains of dairy origin. *Int. J. Food Microbiol.* 91:129–139. [https://doi.org/10.1016/S0168-1605\(03\)00368-4](https://doi.org/10.1016/S0168-1605(03)00368-4).

- Goris, J., K. T. Konstantinidis, J. A. Klappenbach, T. Coenye, P. Vandamme, and J. M. Tiedje. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57:81–91. <https://doi.org/10.1099/ijls.0.64483-0>.
- Hagi, T., M. Kobayashi, and M. Nomura. 2010. Molecular-based analysis of changes in indigenous milk microflora during the grazing period. *Biosci. Biotechnol. Biochem.* 74:484–487. <https://doi.org/10.1271/bbb.90470>.
- Hagi, T., K. Sasaki, H. Aso, and M. Nomura. 2013. Adhesive properties of predominant bacteria in raw cow's milk to bovine mammary gland epithelial cells. *Folia Microbiol. (Praha)* 58:515–522. <https://doi.org/10.1007/s12223-013-0240-z>.
- Hornik, B., J. Czarny, J. Staninska-Pięta, Ł. Wolko, P. Cyplik, and A. Piotrowska-Cyplik. 2021. The raw milk microbiota from semi-subsistence farms characteristics by NGS analysis method. *Molecules* 26:5029. <https://doi.org/10.3390/molecules26165029>.
- Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>.
- Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
- Kimura, K., T. Nishio, C. Mizoguchi, and A. Koizumi. 2010. Analysis of the composition of lactobacilli in humans. *Biosci. Microflora* 29:47–50. <https://doi.org/10.12938/bifidus.29.47>.
- Konstantinidis, K. T., and J. M. Tiedje. 2005. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. USA* 102:2567–2572. <https://doi.org/10.1073/pnas.0409727102>.
- Kudo, Y., K. Oki, and K. Watanabe. 2012. *Lactobacillus delbrueckii* ssp. *sunkii* ssp. nov., isolated from sunki, a traditional Japanese pickle. *Int. J. Syst. Evol. Microbiol.* 62:2643–2649. <https://doi.org/10.1099/ijls.0.037051-0>.
- Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>.
- Li, M., X. Zhou, C. Stanton, R. P. Ross, J. Zhao, H. Zhang, B. Yang, and W. Chen. 2021. Comparative genomics analyses reveal the differences between *B. longum* ssp. *infantis* and *B. longum* ssp. *longum* in carbohydrate utilisation, CRISPR-Cas systems and bacteriocin operons. *Microorganisms* 9:1713. <https://doi.org/10.3390/microorganisms9081713>.
- Liu, W., J. Yu, Z. Sun, Y. Song, X. Wang, H. Wang, T. Wuren, M. Zha, B. Menghe, and Z. Heping. 2016. Relationships between functional genes in *Lactobacillus delbrueckii* ssp. *bulgaricus* isolates and phenotypic characteristics associated with fermentation time and flavor production in yogurt elucidated using multilocus sequence typing. *J. Dairy Sci.* 99:89–103. <https://doi.org/10.3168/jds.2015-10209>.
- Masco, L., M. Ventura, R. Zink, G. Huys, and J. Swings. 2004. Polyphasic taxonomic analysis of *Bifidobacterium animalis* and *Bifidobacterium lactis* reveals relatedness at the subspecies level: Reclassification of *Bifidobacterium animalis* as *Bifidobacterium animalis* subsp. *animalis* subsp. nov. and *Bifidobacterium lactis* as *Bifidobacterium animalis* subsp. *lactis* subsp. nov. *Int. J. Syst. Evol. Microbiol.* 54:1137–1143. <https://doi.org/10.1099/ijls.0.03011-0>.
- Mohammed, M., H. Abd El-Aziz, N. Omran, S. Anwar, S. Awad, and M. El-Soda. 2009. Rep-PCR characterization and biochemical selection of lactic acid bacteria isolated from the Delta area of Egypt. *Int. J. Food Microbiol.* 128:417–423. <https://doi.org/10.1016/j.ijfoodmicro.2008.09.022>.
- Mora, D., M. G. Fortina, C. Parini, G. Ricci, M. Gatti, G. Giraffa, and P. L. Manachini. 2002. Genetic diversity and technological properties of *Streptococcus thermophilus* strains isolated from dairy products. *J. Appl. Microbiol.* 93:278–287. <https://doi.org/10.1046/j.1365-2672.2002.01696.x>.
- Moschetti, G., G. Blaiotta, M. Aponte, P. Catzeddu, F. Villani, P. Deiana, and S. Coppola. 1998. Random amplified polymorphic DNA and amplified ribosomal DNA spacer polymorphism: powerful methods to differentiate *Streptococcus thermophilus* strains. *J. Appl. Microbiol.* 85:25–36. <https://doi.org/10.1046/j.1365-2672.1998.00461.x>.
- Nikoloudaki, O., W. J. F. Lemos Junior, L. Borruso, S. Campanaro, M. De Angelis, R. F. Vogel, R. Di Cagno, and M. Gobbetti. 2021. How multiple farming conditions correlate with the composition of the raw cow's milk lactic microbiome. *Environ. Microbiol.* 23:1702–1716. <https://doi.org/10.1111/1462-2920.15407>.
- Poznanski, E., A. Cavazza, F. Cappa, and P. S. Cocconcelli. 2004. Indigenous raw milk microbiota influences the bacterial development in traditional cheese from an alpine natural park. *Int. J. Food Microbiol.* 92:141–151. <https://doi.org/10.1016/j.ijfoodmicro.2003.09.006>.
- Richter, M., and R. Rosselló-Móra. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. USA* 106:19126–19131. <https://doi.org/10.1073/pnas.0906412106>.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>.
- Shahbal, S., D. Hemme, and M. Desmazeaud. 1991. High cell wall-associated proteinase activity of some *Streptococcus thermophilus* strains (H-strains) correlated with a high acidification rate in milk. *Lait* 71:351–357. <https://doi.org/10.1051/lait:1991327>.
- Sneath, P. H. A., and R. R. Sokal. 1973. Numerical Taxonomy. The Principles and Practice of Numerical Classification. W. H. Freeman and Company. <https://doi.org/10.2307/2412767>.
- Song, Y., Z. Sun, C. Guo, Y. Wu, W. Liu, J. Yu, B. Menghe, R. Yang, and H. Zhang. 2016. Genetic diversity and population structure of *Lactobacillus delbrueckii* subspecies *bulgaricus* isolated from naturally fermented dairy foods. *Sci. Rep.* 6:22704. <https://doi.org/10.1038/srep22704>.
- Stackebrandt, E. 2006. Taxonomic parameters revisited: Tarnished gold standards. *Microbiol. Today* 33:152–155.
- Tamura, K. 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. *Mol. Biol. Evol.* 9:678–687. <https://doi.org/10.1093/oxfordjournals.molbev.a040752>.
- Tanigawa, K., and K. Watanabe. 2011. Multilocus sequence typing reveals a novel subspeciation of *Lactobacillus delbrueckii*. *Microbiology (Reading)* 157:727–738. <https://doi.org/10.1099/mic.0.043240-0>.
- Tanizawa, Y., T. Fujisawa, E. Kaminuma, Y. Nakamura, and M. Arita. 2016. DFAST and DAGA: web-based integrated genome annotation tools and resources. *Biosci. Microbiota Food Health* 35:173–184. <https://doi.org/10.12938/bmfh.16-003>.
- Tanizawa, Y., T. Fujisawa, and Y. Nakamura. 2018. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>.
- Van Hoorde, K., M. Heyndrickx, P. Vandamme, and G. Huys. 2010a. Influence of pasteurization, brining conditions and production environment on the microbiota of artisan Gouda-type cheeses. *Food Microbiol.* 27:425–433. <https://doi.org/10.1016/j.fm.2009.12.001>.
- Van Hoorde, K., I. Van Leuven, P. Dirinck, M. Heyndrickx, K. Cou-dijzer, P. Vandamme, and G. Huys. 2010b. Selection, application and monitoring of *Lactobacillus paracasei* strains as adjunct cultures in the production of Gouda-type cheeses. *Int. J. Food Microbiol.* 144:226–235. <https://doi.org/10.1016/j.ijfoodmicro.2010.05.007>.
- Vandamme, P., B. Pot, M. Gillis, P. de Vos, K. Kersters, and J. Swings. 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol. Rev.* 60:407–438. <https://doi.org/10.1128/mr.60.2.407-438.1996>.
- Weiss, N., U. Schillinger, and O. Kandler. 1983. *Lactobacillus lactis*, *Lactobacillus leichmannii* and *Lactobacillus bulgaricus*, subjective

- synonyms of *Lactobacillus delbrueckii*, and description of *Lactobacillus delbrueckii* ssp. *lactis* comb. nov. and *Lactobacillus delbrueckii* ssp. *bulgaricus* comb. nov. Syst. Appl. Microbiol. 4:552–557. [https://doi.org/10.1016/S0723-2020\(83\)80012-5](https://doi.org/10.1016/S0723-2020(83)80012-5).
- Wijesundera, C., M. Roberts, and G. Limsowtin. 1997. Flavour development in aseptic cheese curd slurries prepared with single-strain starter bacteria in the presence and absence of adjuncts. Lait 77:121–131. <https://doi.org/10.1051/lait:199718>.
- Wouters, J. T. M., E. H. E. Ayad, J. Hugenholtz, and G. Smit. 2002. Microbes from raw milk for fermented dairy products. Int. Dairy J. 12:91–109. [https://doi.org/10.1016/S0958-6946\(01\)00151-0](https://doi.org/10.1016/S0958-6946(01)00151-0).
- Yamamoto, E., R. Watanabe, A. Koizumi, T. Ishida, and K. Kimura. 2020. Isolation and characterization of *Streptococcus thermophilus* possessing *prtS* gene from raw milk in Japan. Biosci. Microbiota Food Health 39:169–174. <https://doi.org/10.12938/bmfh.2019-052>.
- Yoon, S. H., S. M. Ha, J. M. Lim, S. J. Kwon, and J. Chun. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie van Leeuwenhoek 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>.
- Yu, J., Z. Sun, W. Liu, X. Xi, Y. Song, H. Xu, Q. Lv, Q. Bao, B. Menghe, and T. Sun. 2015. Multilocus sequence typing of *Streptococcus thermophilus* from naturally fermented dairy foods in China and Mongolia. BMC Microbiol. 15:236. <https://doi.org/10.1186/s12866-015-0551-0>.
- Zeaiter, Z., P. E. Fournier, H. Ogata, and D. Raoult. 2002. Phylogenetic classification of *Bartonella* species by comparing *groEL* sequences. Int. J. Syst. Evol. Microbiol. 52:165–171. <https://doi.org/10.1099/00207713-52-1-165>.
- Zerbino, D. R., and E. Birney. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
- Zuckermandl, E., and L. Pauling. 1965. Evolutionary divergence and convergence in proteins. Pages 97–166 in *Evolving Genes and Proteins*. V. Bryson and H. J. Vogel, ed. Academic Press.

ORCIDS

H. Tsuchihashi <https://orcid.org/0000-0002-2995-2545>

T. Ishida <https://orcid.org/0000-0002-3669-3932>

K. Kimura <https://orcid.org/0000-0002-7106-2650>