Title	Microbial ecology of human gut bifidobacteria and lactobacilli : their taxonomy and behavior
Author(s)	大木, 海平
Citation	北海道大学. 博士(環境科学) 乙第7108号
Issue Date	2020-12-25
DOI	10.14943/doctoral.r7108
Doc URL	http://hdl.handle.net/2115/80187
Type	theses (doctoral)
File Information	Kaihei_Oki.pdf

Microbial ecology of human gut bifidobacteria and lactobacilli: their taxonomy and behavior

（ヒト腸内 *Bifidobacterium* 属および *Lactobacillus* 属細菌の微生物生態学:その分類と挙動）

Kaihei Oki
2020

Contents

Chapter 1: General introduction

Microbes are distributed in the various environments on Earth and composed of complex communities with responding to environmental variables, e.g. pH, oxygen, nutrients (Thompson, et al. 2017). According to Fenchel (1992), "*The aim of microbial ecology is to find principles which explain the structure and function of microbial communities*". To achieve this goal, it is essential to know "who" reside and "how" it behaves there for each microbe, including its individual function and interaction with other environmental components. The importance of the interaction between host and its symbiont was well recognized recently and they were thought to be considered as a biomolecular networks, i.e. holobiont (Bordenstein and Theis 2015). Human harbors the most complex microbial ecosystem in the gut containing large number of microbiota (10^{10}–10^{11} cells/g feces) encompassing several hundreds of species (Guarner and Malagelada 2003). Healthy adults harbor balanced microbial composition in their gut dominated by the phyla *Firmicutes*, *Bacteroidetes*, *Actinobacteria* and *Proteobacteria* (Cho and Blaser 2012). The imbalance of gut microbiota, i.e. dysbiosis, was reported to associate to various health disorders of their host not only in the gut but also in other body regions. For example, the expansion of *Proteobacteria* including *Escherichia coli*, *Campylobacter concisus* and enterohepatic *Helicobacter* associated with the pathogenesis of inflammatory bowel disease (Mukhopadhya, et al. 2012). The increase of *Proteobacteria* and the decrease of *Firmicutes* were also reported for the patients with advanced fibrosis in nonalcoholic fatty liver disease (Loomba, et al. 2017). As for systemic disorder, an association was reported between the elevation of *Firmicutes* / *Bacteroidetes* ratio and metabolic disease (Ley, et al. 2006). Considering their strong and wide-ranged effect on human health (Postler and Ghosh 2017), we can no longer ignore the influence of human gut microbiota in this field.

The advance in next generation sequencing has expanded our knowledge about human gut microbiota. It has also revealed that they still contain vast number of unassigned microbial components (Lagier, et al. 2012). There are also an approach to estimate the detailed function of human gut microbiota based on functional metagenomics regardless to their culturabilities (Almeida, et al. 2019). Recently, a number of studies have focused on human gut microbiota, especially for understanding its correlation to host's health and diseases (Clemente, et al. 2012; Kho and Lal 2018) as well as to host's properties (e.g. life-stages, genders and geographical regions) (De Filippo, et al. 2010; Flak, et al. 2013; Odamaki, et al. 2016). On the other hand, most of the studies were conducted at genus or higher taxonomic level and the information for the function and behavior of each species or strain was still limited (Lagier, et al. 2016). For the further understanding of the microbial ecosystem in human gut, far more information about the community structure and the behavior needs to be accumulated for each of the human gut microbial components. Accurate classification and identification of human gut microbiota are also essential to estimate their genetic and physiological property. Though recent advance in molecular microbiology resulted in the revision of human gut microbial taxonomy (Duncan, et al. 2002; Liu, et al. 2008; Taras, et al. 2002), there are still remaining taxa which should be reclassified (Haas and Blanchard 2020). Simultaneously, future description of novel taxa would also lead to the necessity to further revision of the current microbial taxonomy. Therefore, it should be improved continuously, along with applying more appropriate classification and identification method.

Among the human gut microbiota, *Bifidobacterium* and *Lactobacillus* have been known as common inhabitants. *Bifidobacterium* is dominant component of human distal gut microbiota especially during early stage of life, and the alteration of their composition was frequently reported in various host's health dysfunction provoked by dysbiosis, e.g. obesity, inflammatory bowel disease and allergy (Tojo, et al. 2014). Although *Lactobacillus* is more minor in distal gut of human, the compositional change of *Lactobacillus* was also reported at this area for several disease, e.g. irritable bowel syndrome, multiple sclerosis and type 1 diabetes (Heeney, et al. 2018; Reuter 2001). Owing to their wide range of beneficial effects to host, e.g. bowel regulation, immunostimulation, intestinal barrier reinforcement, protection from pathogen, some of the strains belonging to these genera were used commercially as probiotics (George Kerry, et al. 2018; Turroni, et al. 2014). This study aims at the further understanding of microbial ecosystem in human gut from the point of views of their taxonomy and behavior, putting focus on these functionally and commercially important components of human gut microbiota.

In chapter 2, I conducted taxonomic studies to propose novel components of human gut lactobacilli and improved subspeciation method for the dominant human gut bifidobacteria.

As the first part (Part 2A), I successfully found and proposed the novel species belonging to the genus *Lactobacillus* originating from human gut. The re-analysis based on 16S rRNA gene sequence similarity was conducted for 328 strains isolated from the feces of 61 healthy Japanese subjects in the different age-groups. As the result, two strains isolated from an adult and a centenarian, respectively, were found to represent novel species. The phenotypic and genotypic features demonstrated that these strains represent independent two novel species of the genus *Lactobacillus*, and the names *Lactobacillus saniviri* sp. nov. and *Lactobacillus senioris* sp. nov. were proposed, respectively.

As the second part (Part 2B), I proposed the rearrangement of the classification of the subspecies of *Bifidobacterium longum*, one of the dominant human and animal gut microbial components. Three molecular biological methods (Amplified Fragment Length Polymorphism [AFLP], Multi Locus Sequence Analysis [MLSA] and Multi Locus Sequence Typing [MLST]) were applied to classify 25 *B. longum* strains previously identified. As the result, the clusters corresponding to three known subspecies (subsp. *longum*, subsp. *infantis* and subsp. *suis*) were observed with each method. Additionally, the cluster of *B. longum* subsp. *suis* was further divided to two subclusters, with or without urease activity. The subcluster without urease activity was proposed as *B. longum* subsp. *suillum* named from the origin, porcine feces, as the fourth subspecies of *B. longum*.

In chapter 3, I conducted a study on a behavior of bifidobacterial strains colonized in infant’s gut where bifidobacteria became the most dominant component of human gut microbiota.

As the part 3A, I investigated whether a strain colonized in an individual’s gut continued to exist there over the long term from early infancy, with focusing on *B. longum* subsp. *longum*, a common component of infant gut microbiota. The 462 strains were isolated over time from the feces of 12 subjects followed from early infancy (the first six months of life) up to childhood (approximately six years of age) as well as from their mothers' perinatal samples (prenatal feces and postnatal breast milk). As the result of genotype comparison, several strains were confirmed to colonize and persist in single individuals from early infancy through more than six years, one of which was also detected from the corresponding mother's postnatal breast milk, confirming the existence of long-term colonizers from this period. Moreover, the results suggested that these strains persisted in the subjects' gut while co-existing with the other predominant bifidobacterial species.

References

1. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et al. 2019. A new genomic blueprint of the human gut microbiota. Nature **568**:499–504.
2. Bordenstein SR, Theis KR. 2015. Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. PLoS Biol **13**:1–23.
3. Cho I, Blaser MJ. 2012. The human microbiome: At the interface of health and disease. Nat Rev Genet. NIH Public Access.
4. Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. Cell **148**:1258–1270.
5. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, et al. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A **107**:14691–14696.
6. Duncan SH, Hold GL, Harmsen HJM, Stewart CS, Flint HJ. 2002. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int J Syst Evol Microbiol **52**:2141–2146.
7. Fenchel T. 1992. What can ecologists learn from microbes: life beneath a square centimetre of sediment surface. Funct Ecol **6**:499–507.
8. Flak MB, Neves JF, Blumberg RS. 2013. Immunology. Welcome to the microgenderome. Science **339**:1044–1045.
9. George Kerry R, Patra JK, Gouda S, Park Y, Shin HS, Das G. 2018. Benefaction of

probiotics for human health: A review. J Food Drug Anal **26**:927–939.

10. Guarner F, Malagelada JR. 2003. Gut flora in health and disease. Lancet **361**:512–519.
11. Haas KN, Blanchard JL. 2020. Reclassification of the *Clostridium clostridioforme* and *Clostridium sphenoides* clades as *Enterocloster* gen. Nov. and *Lacrimispora* gen. nov., including reclassification of 15 taxa. Int J Syst Evol Microbiol **70**:23–34.
12. Heeney DD, Gareau MG, Marco ML. 2018. Intestinal *Lactobacillus* in health and disease, a driver or just along for the ride? Curr Opin Biotechnol. Elsevier Ltd.
13. Kho ZY, Lal SK. 2018. The human gut microbiome - A potential controller of wellness and disease. Front Microbiol **9**:1835.
14. Lagier J-C, Million M, Hugon P, Armougom F, Raoult D. 2012. Human gut microbiota: repertoire and variations. Front Cell Infect Microbiol **2**:1–19.
15. Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. 2016. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol **1**:1–8.
16. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. 2006. Microbial ecology: human gut microbes associated with obesity. Nature **444**:1022–1023.
17. Liu C, Finegold SM, Song Y, Lawson PA. 2008. Reclassification of *Clostridium coccoides*, *Ruminococcus hansenii*, *Ruminococcus hydrogenotrophicus*, *Ruminococcus luti*, *Ruminococcus productus* and *Ruminococcus schinkii* as *Blautia coccoides* gen. nov., comb. nov., *Blautia hansenii* comb. nov., *Blautia hydrogenotrophica* comb. nov., *Blautia luti* comb. nov., *Blautia producta* comb. nov., *Blautia schinkii* comb. nov. and description of *Blautia wexlerae* sp. nov., isolated from human faeces. Int J Syst Evol Microbiol **58**:1896–1902.
18. Loomba R, Seguritan V, Li W, Long T, Klitgord N, Bhatt A, et al. 2017. Gut microbiome-based metagenomic signature for non-invasive detection of advanced fibrosis in human nonalcoholic fatty liver disease. Cell Metab **25**:1054–1062.
19. Mukhopadhya I, Hansen R, El-Omar EM, Hold GL. 2012. IBD-what role do *Proteobacteria* play? Nat Rev Gastroenterol Hepatol. Nature Publishing Group.

20. Odamaki T, Kato K, Sugahara H, Hashikura N, Takahashi S, Xiao JZ, et al. 2016. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. BMC Microbiol **16**:90.

21. Postler TS, Ghosh S. 2017. Understanding the holobiont: how microbial metabolites affect human health and shape the immune system. Cell Metab. Cell Press.

22. Reuter G. 2001. The *Lactobacillus* and *Bifidobacterium* microflora of the human intestine: composition and succession. Curr Issues Intest Microbiol **2**:43–53.

23. Taras D, Simmering R, Collins MD, Lawson PA, Blaut M. 2002. Reclassification of *Eubacterium formicigenerans* Holdeman and Moore 1974 as *Dorea formicigenerans* gen. nov., comb. nov., and description of *Dorea longicatena* sp. nov., isolated from human faeces. Int J Syst Evol Microbiol **52**:423–428.

24. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. Nature **551**:457–463.

25. Tojo R, Suárez A, Clemente MG, De Los Reyes-Gavilán CG, Margolles A, Gueimonde M, et al. 2014. Intestinal microbiota in health and disease: Role of bifidobacteria in gut homeostasis. World J Gastroenterol. WJG Press.

26. Turroni F, Ventura M, Buttó LF, Duranti S, O'Toole PW, Motherway MOC, et al. 2014. Molecular dialogue between the human gut microbiota and the host: a *Lactobacillus* and *Bifidobacterium* perspective. Cell Mol Life Sci **71**:183–203.

Chapter 2: Taxonomic studies on human gut *Bifidobacterium* and *Lactobacillus*

Part 2A

Lactobacillus saniviri sp. nov. and *Lactobacillus senioris* sp. nov., isolated from human feces

2A.1. Introduction

Lactobacillus strains are one of the most familiar microorganisms to humans, not only because of their association with a wide range of naturally fermented dairy products, grain crops and vegetables but also because they inhabit the gastrointestinal tracts of humans and animals, and are the most widely used probiotics aimed at promoting a healthy lifestyle. At the time of this study (in 2011), the genus *Lactobacillus* comprises about 140 species (currently, 253 species and 23 subspecies were described as correct name), although only a few species – *Lactobacillus brevis*, the *Lactobacillus casei* group, *Lactobacillus delbrueckii*, *Lactobacillus gasseri*, *Lactobacillus parabuchneri*, the *Lactobacillus plantarum* group, *Lactobacillus reuteri* and *Lactobacillus sakei* – have been found in human feces (Bello, et al. 2003; Walter, et al. 2001). Although the names of around 80 species have been validly published in the first decade of 2000, no novel species isolated from human feces have been described during this time (J.P. Euzéby: List of Prokaryotic names with Standing in Nomenclature; https://www.bacterio.net/). In this study, we explored the strains comprised novel *Lactobacillus* species isolated from human feces.

2A.2. Materials and methods

2A.2.1. Bacterial strains used in this study

In order to explore novel *Lactobacillus* strains, we reanalyzed the culture collection of the Yakult Central Institute. The following *Lactobacillus* strains were obtained from the culture collection of Yakult Central Institute (YIT, Tokyo, Japan) and were used as references: *L. casei* YIT 0078^T, *Lactobacillus paracasei* subsp. *paracasei* YIT 0209^T and *Lactobacillus rhamnosus* YIT 0105^T (for isolate YIT 12363^T); and *Lactobacillus parafarraginis* YIT 12274^T and *Lactobacillus rapi* YIT 11204^T (for isolate YIT 12364^T).

2A.2.2. Growth condition and DNA extraction

The strains used in this study were cultivated and maintained in MRS broth (BD Difco, MA, USA) (De Man, et al. 1960) at 37°C (or 30°C for *L. parafarraginis* YIT 12274^T) for 16 h, unless indicated otherwise. Modified MRS broth [MRS broth supplemented with 1% (w/v) L-arabinose and 1% (w/v) maltose instead of glucose] was used for *L. rapi* YIT 11204^T. Chromosomal DNA used as a template for 16S rRNA, *rpoA* and *pheS* gene sequence amplification was prepared according to the method of Watanabe (2008).

2A.2.3. Sequencing and phylogenetic analysis

The conditions for PCR amplification of the partial 16S rRNA gene and subsequent DNA sequencing have been described previously (Chao, et al. 2008). The *pheS* and *rpoA* gene sequences for YIT 12363^T, YIT 12364^T, *Lactobacillus buchneri* YIT 0077^T (*rpoA* gene only), *Lactobacillus camelliae* YIT 12276^T, *Lactobacillus diolivorans* YIT 10368^T (*rpoA* gene only), *Lactobacillus farraginis* YIT 12273^T, *Lactobacillus kisonensis* YIT 11168^T, *Lactobacillus otakiensis* YIT 11163^T, *L. parafarraginis* YIT 12274^T, *L. rapi* YIT 11204^T and *Lactobacillus sunkii* YIT 11161^T were amplified by PCR with the primer set of rpoA-21F (5´-ATGATYGARTTTGAAAAAACC-3′) and rpoA-23R (5′-ACHGTRTTRATDCCDGCRCG-3′) (Naser, et al. 2005) as well as pheS-21F (5′-CAYCCNGCHCGYGAYATGC-3′) and pheS-952R (5′-TATTTCAAAATTGCRAAACGRTC-3′) (for YIT 12363^T, YIT 12364^T) or pheS-23R (5′-GGRTGTACCATVCCNGCHCC-3′) (Naser, et al. 2005), respectively. The PCR mixture (25 µl) contained 1×Ex *Taq* Buffer, 200 µM of each dNTP, 0.4 µM of each primer, 1 U Ex *Taq* HS polymerase (Takara Bio) and 10 ng template DNA. The amplification program consisted of: one cycle of 95°C for 2 min; 35 cycles of 95°C for 20 s, 55°C (for *pheS* gene) or 46°C (for *rpoA* gene) for 30 s, 72°C for 1 min 15 s; and finally one cycle of 72°C for 3 min. For amplification of the *pheS* gene of strains YIT 12363T and YIT 12364T, 1×*Taq* buffer and *Taq* polymerase (Takara Bio) were used and the annealing temperature was changed to 42°C. For amplification of the *pheS* gene of *L. rapi* YIT 11204^T, the annealing temperature was changed to 42°C. For amplification of the *rpoA* gene of *L. farraginis* YIT 12274^T, the annealing and elongation steps were modified to 48°C for 20 s and 72°C for 5 s, respectively.

The amplified 16S rRNA, *pheS* and *rpoA* genes were purified by using an AMPure® kit (Beckman Coulter, Brea, CA, USA) and were subsequently sequenced using the ABI PRISM BigDye® Terminator v3.1 cycle sequencing kit (Applied Biosystems, CA, USA). The closest known species to the novel isolates were determined by FASTA and sequences of members of the most closely related species were extracted from GenBank/EMBL/DDBJ.

Multiple alignment and the construction of phylogenetic trees were performed with the program CLUSTAL_X v. 2.0.12. (Thompson, et al. 1997). For strain YIT 12363^T, strain YIT 12364^T and related strains, approximately 1390 bp of the 16S rRNA gene, approximately 280 bp of the *pheS* gene and approximately 350 bp of the *rpoA* gene sequences were used for constructing phylogenetic trees by the neighbor-joining method (Saitou and Nei 1987). The statistical reliability of trees was evaluated by bootstrap analysis of 1000 replicates (Felsenstein 1985) and tree topologies were also confirmed with the minimum-evolution (Rzhetsky and Nei 1992) and maximum-parsimony (Fitch 1971) methods, by using MEGA v4.1 software (Tamura, et al. 2007), according to the Kimura two-parameter model (Kimura 1980).

2A.2.4. DNA-DNA hybridization and G+C content

For determination of DNA–DNA relatedness and the DNA G+C content, chromosomal DNA was extracted according to the method of Marmur (1961). DNA–DNA hybridization analyses were performed between strain YIT 12363^T, *L. casei* YIT 0078^T, *L. paracasei* subsp. *paracasei* YIT 0209^T and *L. rhamnosus* YIT 0105^T and between strain YIT 12364^T, *L. parafarraginis* YIT 12274^T and *L. rapi* YIT 11204^T, respectively. The microdilution well technique was used as described by Ezaki, et al. (1989) using a Spectra Max M2 (Molecular Device, CA, USA) for fluorescence measurements. Reciprocal hybridization experiments were performed for every pair of strains at 44°C (for YIT 12363^T and its reference strains) or 40°C (for YIT 12364^T and its reference strains) for 2 h in the presence of 50% formamide, using biotinylated DNA and unlabeled ssDNA, which was bound non-covalently to microplate wells. The highest and lowest values of eight replicate wells were excluded and the mean of the remaining values was calculated for each experiment.

To assess the DNA G+C content, DNA was enzymatically degraded into nucleosides as described previously (Mesbah, et al. 1989) and then separated by HPLC.

2A.2.5. Morphological, cultural and biochemical tests

The following tests were performed according to standard techniques at 37°C unless otherwise stated. Cell shape, cell size and Gram staining were determined by using cultures grown in MRS broth at 37°C for 16 h. Motility was tested in MRS soft agar (0.15%). Catalase activity was determined by using cells grown on MRS agar. Gas production from glucose was measured with a Durham tube in MRS broth. Production of dextran was assessed on MRS agar in which glucose was replaced with 2% (w/v) sucrose. The methods of Barrow and Feltham (1993) were used to determine growth at various temperatures and pH and in the presence of NaCl, reduction of nitrate and production of ammonia from arginine. Carbohydrate fermentation tests were conducted by using the API 50 CHL system (bioMérieux, Marcy-l'Étoile, France) according to the manufacturer's instructions. The presence of diaminopimelic acid in the cell-wall peptidoglycan was determined according to Kandler and Weiss (1986). Cell-wall peptidoglycan for analysis of amino acid composition was prepared and hydrolyzed according to the method of Schleifer and Kandler (1972). Cell-wall amino acids were analyzed by HPLC (Alliance 2695 HPLC system; Waters, Milford, MA, USA) equipped with a fluorescence detector (model 474 Fluorescence Detector; Waters) and an AccQ-Tag column (3.9×150 mm; Waters), by using an AccQ-Fluor reagent KIT (6-aminoquinolyl-*N*-hydroxysuccinimidylcarbamate; Waters) for derivatization. Cellular fatty acid methyl esters were obtained from cells grown in MRS broth at 37°C (or 30°C for *L. rapi* YIT 11204^T) for 18 h by saponification, methylation and extraction using the method of Miller (1982) with minor modifications (Kuykendall, et al. 1988). FAMEs were determined by using the MIDI system with MOORE5 of the MIS Standard Libraries. The isomers of lactic acid formed from glucose were determined with a TC D-/L-lactic acid test kit (Boehringer-Mannheim, Mannheim, Germany).

2A.3. Results and Discussion

2A.3.1 Phylogenetic characteristics of novel strains

Form the survey of 328 strains isolated from the feces of 61 healthy Japanese subjects in different age-groups – neonates, infants, adults, elderly people and centenarians –, and the result of reanalysis based on 16S rRNA gene sequence similarity values, it was revealed that, YIT 12363^T and YIT 12364^T isolated from the feces of an adult and a centenarian could not be clearly placed within any recognized species of the genus *Lactobacillus*.

Phylogenetic analysis based on 16S rRNA gene sequences showed that isolate YIT 12363^T was a close neighbor to the *L. casei* group of lactobacilli (Fig. 1). 16S rRNA gene sequence similarities between strain YIT 12363^T and its closest neighbors, *L. casei* YIT 0078^T, *L. paracasei* subsp. *paracasei* YIT 0209^T, *L. paracasei* subsp. *tolerans* NBRC 15906^T and *L. rhamnosus* YIT 0105^T were high, with values of 95.3, 95.6, 95.3 and 95.4%, respectively. Similar topologies were obtained by the minimum-evolution (Supplementary Fig. S1) and maximum-parsimony (data not shown) methods. Similarities between the *pheS* and *rpoA* gene sequences of strain YIT 12363^T and the most closely related strains ranged from 73.5 to 75.2% and from 79.5 to 80.4%, respectively (Supplementary Figs S2 and S3). Phylogenetic analysis based on 16S rRNA gene sequences placed isolate YIT 12364^T in the *L. buchneri* group of lactobacilli (Fig. 1). 16S rRNA gene sequence similarities between strain YIT 12364^T and its closest neighbors, *L. diolivorans* YIT 10368^T, *L. parafarraginis* NRIC 0677^T and *L. rapi* YIT 11204^T, were 95.8, 96.0 and 96.0%, respectively; 16S rRNA gene sequence similarities with the type strains of all other species in this group were below 97%. Similar topologies were obtained by the minimum-evolution (Supplementary Fig. S1) and maximum-parsimony (data not shown) methods. Similarities between the *pheS* and *rpoA* gene sequences of isolate YIT 12364^T and the most closely related strains ranged from 77.3 to 79.9% and from 79.9 to 80.0%, respectively (Supplementary Figs S2 and S3). Interspecies gaps within the genus *Lactobacillus* based on *pheS* and *rpoA* gene sequences normally exceed 10% and 5%, respectively (Naser, et al. 2007) , which suggested that the two identified strains should be classified as members of novel species.

The levels of DNA–DNA relatedness of strain YIT 12363^T to *L. casei* YIT 0078^T, *L.*

paracasei subsp. *paracasei* YIT 0209^T and *L. rhamnosus* YIT 0105^T were 22.3–30.0, 19.7–25.7 and 12.5–27.8%, respectively; those of strain YIT 12364^T to *L. parafarraginis* YIT 12274^T and *L. rapi* YIT 11204^T were 11.6–26.3 and 15.9–23.1%, respectively. All values among the reference strains were well below the 70% cut-off value that indicates separate species status (Stackebrandt and Goebel 1994). The DNA G+C contents of strains YIT 12363^T and YIT 12364^T were 48.1 and 37.9 mol%, respectively, which are within the range for the genus *Lactobacillus* (32–59 mol%) (Salvetti, et al. 2012).

2A.3.2. Phylogenetic characteristics of novel strains

Various phenotypic characteristics of isolate YIT 12363^T differed from those of *L. casei* YIT 0078^T, *L. paracasei* subsp. *paracasei* YIT 0209^T and *L. rhamnosus* YIT 0105^T. Likewise, isolate YIT 12364^T had different phenotypic characteristics from *L. parafarraginis* YIT 12274^T and *L. rapi* YIT 11204^T (Tables 1, 2 and 3). According to the genotypic and phenotypic data obtained, it is proposed that the identified strains should be classified as representatives of two novel *Lactobacillus* species: *L. saniviri* sp. nov. (YIT 12363^T) and *L. senioris* sp. nov. (YIT 12364^T).

2A.4. Description of novel species

2A.4.1. *Lactobacillus saniviri* sp. nov.

Lactobacillus saniviri (sa.ni.vi′ri. L. adj. *sanus* healthy; L. gen. n. *viri* of an adult male; N.L. gen. n. *saniviri* of a healthy adult male, indicating the source of the type strain).

Cells are rod-shaped (1.0×1.5–3.0 µm) and occur singly, in pairs or in short chains comprising three to six cells. Cells are Gram-stain-positive, non-motile, asporogenous and facultatively anaerobic. Catalase and pseudocatalase are not produced. After anaerobic growth at 37°C for 72 h, colonies on MRS agar are circular to slightly irregular, 1.5–2.0 mm in diameter and beige with a smooth surface. After aerobic growth at 37°C for 72 h on blood agar, α-hemolysis is observed. In MRS broth, growth occurs at 10°C but not at 45°C. Growth occurs at pH 4.0 and pH 8.5. Growth occurs in the presence of 5% NaCl but not in the presence of 8% NaCl. Gas is not produced from glucose. Both L- (49%) and D-lactate (51%) are produced as end products from glucose. Ammonia is not produced from arginine. Nitrate is not reduced. Acid is produced from glycerol, D-ribose, D-galactose, D-glucose, D-fructose, D-mannose, D-mannitol, *N*-acetylglucosamine, amygdalin, arbutin, salicin, cellobiose, maltose, melibiose, sucrose, trehalose, melezitose (weakly), raffinose, starch (weakly), gentiobiose, D-tagatose, gluconate (weakly) and 2-ketogluconate. Aesculin is hydrolyzed. Dextran is not produced from sucrose. Cells do not contain *meso*-diaminopimelic acid in their cell-wall peptidoglycan. Peptidoglycan structure is of the L-Lys–D-Asp type in the presence of Lys, Glu, Ala and Asp. The major cellular fatty acids (>10%) are unsaturated fatty acid $C_{18:1}\omega 9c$ and cyclopropane C_{19} cyc 9,10. Phylogenetic analysis of the 16S rRNA gene sequence indicates that members of the *L. casei* group of lactobacilli are the closest neighbors.

The type strain, YIT 12363^T (=JCM 17471^T=DSM 24301^T), was isolated from feces of a healthy Japanese adult male in Tokyo, Japan. The DNA G+C content of the type strain is 48.1 mol%.

2A.4.2. *Lactobacillus senioris* sp. nov.

Lactobacillus senioris (se.ni′o.ris. L. gen. n. *senioris* of/from an elderly person, indicating the source of the type strain).

Cells are rod-shaped (0.7×1.0–10.0 μm) and occur singly, in pairs or in chains comprising three to four cells. Cells are Gram-stain-positive, non-motile, asporogenous and facultatively anaerobic. Catalase and pseudocatalase are not produced. After anaerobic growth at 37°C for 72 h, colonies on MRS agar are circular, 1–2 mm in diameter and beige with a smooth or rough surface. In MRS broth, growth occurs at 15°C but not at 10°C or 45°C. Growth does not occur at pH 4.0 or pH 8.5. Growth occurs weakly in the presence of 5% NaCl but not in the presence of 8% NaCl. Gas is produced from glucose. Both L- (72%) and D-lactate (28%) are produced as the end products from glucose. Ammonia is produced from arginine. Nitrate is not reduced. Acid is produced from L-arabinose, D-ribose, D-xylose, D-glucose, D-fructose, *N*-acetylglucosamine (weakly) and gluconate. Aesculin is not hydrolyzed. Dextran is not produced from sucrose. Cells do not contain *meso*-diaminopimelic acid in their cell-wall peptidoglycan. Peptidoglycan structure is of the L-Lys–D-Asp type in the presence of Lys, Glu, Ala and Asp. The major cellular fatty acids are unsaturated fatty acid $C_{18:1}\omega9c$ and cyclopropane C_{19} cyc 9,10. Phylogenetic analysis of the 16S rRNA gene sequence places the species in the *L. buchneri* group of lactobacilli.

The type strain, YIT 12364^T (=JCM 17472^T=DSM 24302^T), was isolated from feces of a 100-year-old elderly female person in Okinawa, Japan. The DNA G+C content of the type strain is 37.9 mol%.

References

1. Barrow GI, Feltham RK, eds. 1993. Cowan and Steel's manual for the identification of medical bacteria, 3rd edn. Cambridge University Press, London.
2. Bello FD, Walter J, Hammes W, Hertel C. 2003. Increased complexity of the species composition of lactic acid bacteria in human feces revealed by alternative incubation condition. Microb Ecol **45**:455–463.
3. Chao SH, Tomii Y, Watanabe K, Tsai YC. 2008. Diversity of lactic acid bacteria in fermented brines used to make stinky tofu. Int J Food Microbiol **123**:134–141.
4. De Man JC, Rogosa M, Sharpe ME. 1960. A medium for the cultivation of lactobacilli. J Appl Bacteriol **23**:130–135.
5. Ezaki T, Hashimoto Y, Yabuuchi E. 1989. Fluorometric deoxyribonucleic acid-deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane filter hybridization in which radioisotopes are used to determine genetic relatedness among bacterial strains. Int J Syst Bacteriol **39**:224–229.
6. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution (N Y) **39**:783–791.
7. Fitch WM. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst Zool **20**:406.
8. Kandler O, Weiss N. 1986. Genus *Lactobacillus* Beijerinck 1901, 212^{AL}. pp. 1209–1234. In: Sneath PHA, Mair NS, Sharpe ME, Holt JG (Eds.), Bergey's Manual of Systematic

Bacteriology vol 2. Williams & Wilkin, Baltimore, MD.

9. Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol **16**:111–120.

10. Kuykendall LD, Roy MA, O'Neill JJ, Devine TE. 1988. Fatty acids, antibiotic resistance, and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum*. Int J Syst Bacteriol **38**:358–361.

11. Marmur J. 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J Mol Biol **3**:208-IN1.

12. Mesbah M, Premachandran U, Whitman WB. 1989. Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. Int J Syst Bacteriol **39**:159–167.

13. Miller LT. 1982. Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids. J Clin Microbiol **16**:584–586.

14. Naser SM, Dawyndt P, Hoste B, Gevers D, Vandemeulebroecke K, Cleenwerck I, et al. 2007. Identification of lactobacilli by *pheS* and *rpoA* gene sequence analyses. Int J Syst Evol Microbiol **57**:2777–2789.

15. Naser SM, Thompson FL, Hoste B, Gevers D, Dawyndt P, Vancanneyt M, et al. 2005. Application of multilocus sequence analysis (MLSA) for rapid identification of *Enterococcus* species based on *rpoA* and *pheS* genes. Microbiology **151**:2141–2150.

16. Rzhetsky A, Nei M. 1992. A simple method for estimating and testing minimum-evolution trees. Mol Biol Evol **9**:945–967.

17. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4**:406–425.

18. Salvetti E, Torriani S, Felis GE. 2012. The genus *Lactobacillus*: a taxonomic update. Probiotics Antimicrob Proteins **4**:217–226.

19. Schleifer KH, Kandler O. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev **36**:407–477.

20. Stackebrandt E, Goebel B. 1994. Taxonomic note: a place for DNA-DNA reassociation and

16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Bacteriol **44**:846–849.

21. Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol **24**:1596–1599.

22. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25**:4876–4882.

23. Walter J, Hertel C, Tannock GW, Lis CM, Munro K, Hammes WP. 2001. Detection of *Lactobacillus*, *Pediococcus*, *Leuconostoc*, and *Weissella* species in human feces by using group-specific PCR primers and denaturing gradient gel electrophoresis. Appl Environ Microbiol **67**:2578–2585.

24. Watanabe K, Fujimoto J, Sasamoto M, Dugersuren J, Tumursuh T, Demberel S. 2008. Diversity of lactic acid bacteria and yeasts in Airag and Tarag, traditional fermented milk products of Mongolia. World J Microbiol Biotechnol **24**:1313–1325.

Tables and Figures

Table 1. Differential characteristics of *L. saniviri* sp. nov. YIT 12363^{T} and closely related lactobacilli.

Characteristic	1	2	3	4
Growth at:				
10 °C	+	w	w	w
45 °C	−	−	−	+
pH 8.5	+	+	−	+
Tolerance to 5 % salt	+	+	−	+
Acid production from :				
Glycerol	+	−	−	−
D-Ribose	+	−	+	+
L-Sorbose	−	−	−	+
L-Rhamnose	−	−	−	w
D-Sorbitol	−	−	−	+
Methyl α-D-glucopyranoside	−	−	w	+
Arbutin	+	+	+	w
Salicin	+	+	+	w
D-Lactose	−	+	+	+
D-Melibiose	+	−	−	−
D-Sucrose	+	w	+	+
D-Melezitose	+	+	+	w
D-Raffinose	+	−	−	−
Starch	w	−	−	−
Gentiobiose	+	+	+	w
D-Turanose	−	−	+	+
D-Lyxose	−	−	−	+
2-Ketogluconate	w	−	−	−
Optical form of lactic acid	DL	L	L	L
Peptidoglycan				
Type	L-Lys-D-Asp	ND	L-Lys-D-Asp	ND
meso -DAP	−	+	−	−
DNA G + C content (mol%)	48.1	48.7	47.5	46.8

Strains: 1, *L. saniviri* sp. nov. YIT 12363^{T}; 2, *L. casei* YIT 0078^{T}; 3, *L. paracasei* subsp. *paracasei* YIT 0209^{T}; 4, *L. rhamnosus* YIT 0105^{T}.

All strains were able to ferment D-galactose, D-glucose, D-fructose, D-mannose, D-mannitol, *N*-acetylglucosamine, amygdalin, cellobiose, maltose, trehalose, D-tagatose and gluconate (weakly). All strains were able to grow at pH 4.0 and 15°C and hydrolyze aesculin. No strains were able to ferment erythritol, D- or L-arabinose, D- or L-xylose, D-adonitol, methyl β- D-xylopyranoside, dulcitol, inositol, methyl α-D-mannopyranoside, inulin, glycogen, xylitol, D- or L-fucose, D- or L-arabitol, or 5-ketogluconate. None of the strains grew in 8% NaCl. Strains were not motile and did not produce gas from glucose, dextran from sucrose, ammonia from arginine, or reduce nitrate. +, positive. −, negative. W, weakly positive reaction. ND, no data. DAP, Diaminopimelic acid.

Table 2. Differential characteristics of *L. senioris* sp. nov. YIT 12364^T and closely related lactobacilli.

Characteristic	1	2	3
Growth at:			
15 °C	+	−	+
pH 4.0	−	+	+
Tolerance to 5 % salt	w	−	−
Ammonia production from arginine	+	+	−
Acid production from :			
Methyl β-D-xylopyranoside	−	+	−
D-Galactose	−	−	w
D-Fructose	+	w	+
Methyl α-D-glucopyranoside	−	+	−
N-Acethylglucosamine	w	−	−
D-Maltose	−	+	+
D-Lactose	−	−	w
D-Melibiose	−	+	+
D-Sucrose	−	w	+
D-Melezitose	−	+	+
D-Raffinose	−	+	+
D-Turanose	−	+	−
Gluconate	+	w	w
5-Ketogluconate	−	w	w
Optical form of lactic acid	DL	L	DL
Peptidoglycan Type	L-Lys-D-Asp	L-Lys-D-Asp	ND
DNA G + C content (mol%)	37.9	41.6	42

Strains: 1. *L. senioris* sp. nov. YIT 12364^T; 2. *L. rapi* YIT 11204^T; 3. *L. parafarraginis* YIT 12274^T.

All strains were able to ferment L-arabinose, D-ribose, D-xylose and D-glucose. No strains were able to ferment glycerol, erythritol, D-arabinose, L-xylose, D-adonitol, D-mannose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl α-D-mannopyranoside, amygdalin, arbutin, salicin, cellobiose, trehalose, inulin, starch, glycogen, xylitol, gentiobiose, D-lyxose, D-tagatose, D- or L-fucose, D- or L-arabitol or 2-ketogluconate. None of the strains grew in 8% NaCl or at 10°C, 45°C and pH 8.5. Strains were not motile, did not contain *meso*-diaminopimelic acid in their peptidoglycan, and were not able to produce dextran from sucrose, reduce nitrate or hydrolyze aesculin. +: positive. −, negative. W, weakly positive reaction. ND, no data.

Table 3. Cellular fatty acid compositions of the novel strains and genetically closely related lactobacilli.

Fatty acid	1	2	3	4
Saturated				
$C_{14:0}$	1.16	11.98	—	1.22
$C_{16:0}$	4.1	17.88	2.23	18.2
$C_{16:0}$ 3OH	—	0.36	—	—
$C_{18:0}$	0.82	0.56	1.53	0.59
$C_{18:0}$ 12OH	4.42	4.74	—	—
Unsaturated				
$C_{16:1}\omega 5c$	—	0.56	—	—
$C_{16:1}\omega 7c$	—	4.63	—	2.28
$C_{18:1}\omega 9c$	56.61	30.32	57.22	28.18
$C_{18:1}\omega 7c$ DMA	0.71	0.76	0.57	—
$C_{18:2}\omega 6,9c$	—	—	1.09	—
Cyclopropane				
C_{19} cyc 9,10	29.29	18.66	34.39	14.9
C_{19} cyc 11,12	—	—	—	5.41
Summed features[a]				
Summed features 10	2.3	7.94	2.05	29.22
Summed features 12	—	0.54	0.93	—
Unknown fatty acid (ECL 18.199)	0.59	1.05	—	—

Strains: 1, *L. saniviri* sp. nov. YIT 12363^T; 2, *L. paracasei* subsp. *paracasei* YIT 0209^T; 3, *L. senioris* sp. nov. YIT 12364^T; 4, *L. rapi* YIT 11204^T.

Values are percentages of total fatty acids.

[a] Summed feature 10, $C_{18:1}\omega 7c$ and/or unknown fatty acid (ECL 17.834). Summed feature 12, unknown fatty acid (ECL 18.622) and/or iso-$C_{19:0}$.

−, not detected. DMA, dimethyl acetal. ECL, equivalent chain-length.

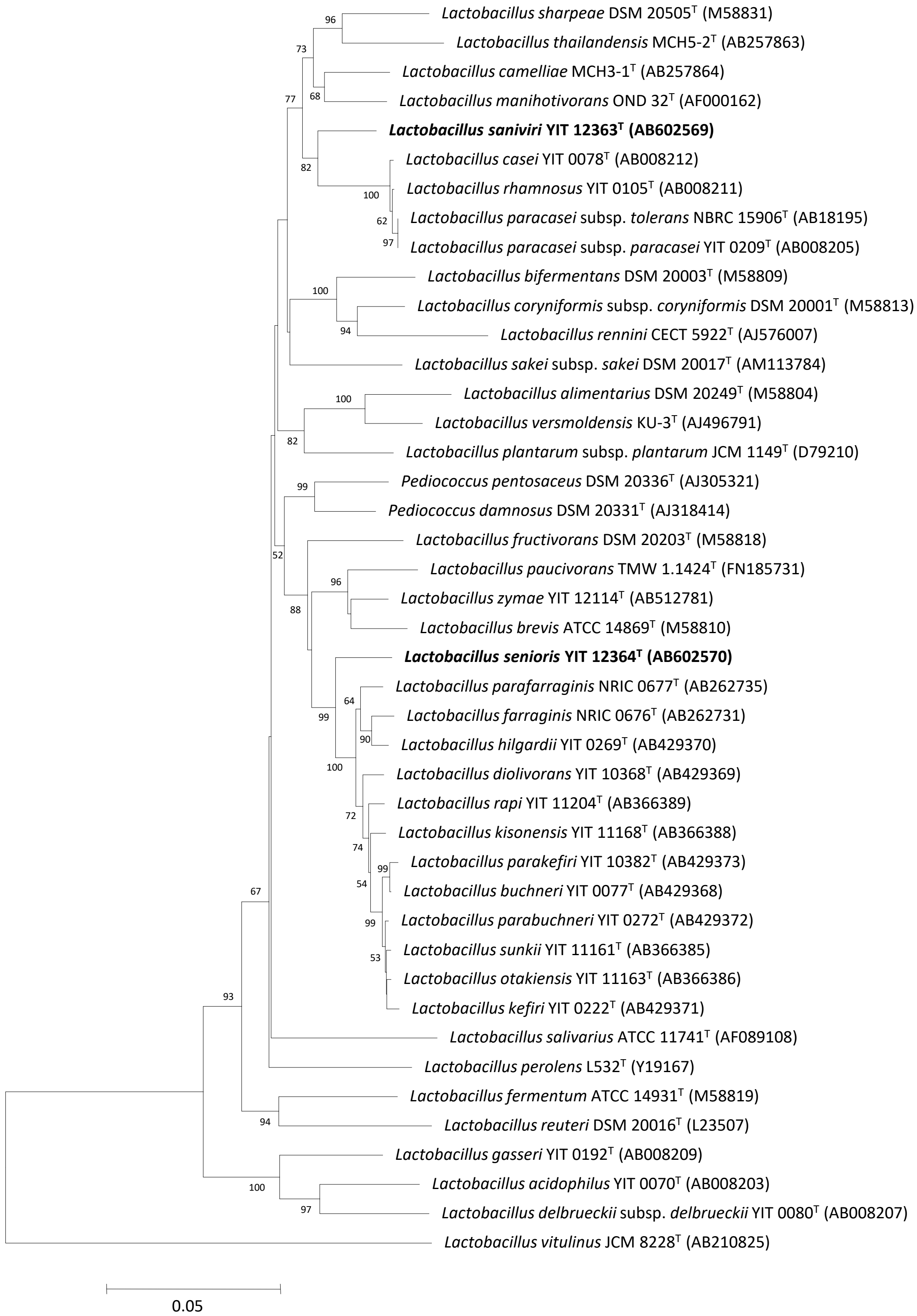

Lactobacillus sharpeae DSM 20505^{T} (M58831)
Lactobacillus thailandensis MCH5-2^{T} (AB257863)
Lactobacillus camelliae MCH3-1^{T} (AB257864)
Lactobacillus manihotivorans OND 32^{T} (AF000162)
Lactobacillus saniviri YIT 12363^{T} (AB602569)
Lactobacillus casei YIT 0078^{T} (AB008212)
Lactobacillus rhamnosus YIT 0105^{T} (AB008211)
Lactobacillus paracasei subsp. *tolerans* NBRC 15906^{T} (AB18195)
Lactobacillus paracasei subsp. *paracasei* YIT 0209^{T} (AB008205)
Lactobacillus bifermentans DSM 20003^{T} (M58809)
Lactobacillus coryniformis subsp. *coryniformis* DSM 20001^{T} (M58813)
Lactobacillus rennini CECT 5922^{T} (AJ576007)
Lactobacillus sakei subsp. *sakei* DSM 20017^{T} (AM113784)
Lactobacillus alimentarius DSM 20249^{T} (M58804)
Lactobacillus versmoldensis KU-3^{T} (AJ496791)
Lactobacillus plantarum subsp. *plantarum* JCM 1149^{T} (D79210)
Pediococcus pentosaceus DSM 20336^{T} (AJ305321)
Pediococcus damnosus DSM 20331^{T} (AJ318414)
Lactobacillus fructivorans DSM 20203^{T} (M58818)
Lactobacillus paucivorans TMW 1.1424^{T} (FN185731)
Lactobacillus zymae YIT 12114^{T} (AB512781)
Lactobacillus brevis ATCC 14869^{T} (M58810)
Lactobacillus senioris YIT 12364^{T} (AB602570)
Lactobacillus parafarraginis NRIC 0677^{T} (AB262735)
Lactobacillus farraginis NRIC 0676^{T} (AB262731)
Lactobacillus hilgardii YIT 0269^{T} (AB429370)
Lactobacillus diolivorans YIT 10368^{T} (AB429369)
Lactobacillus rapi YIT 11204^{T} (AB366389)
Lactobacillus kisonensis YIT 11168^{T} (AB366388)
Lactobacillus parakefiri YIT 10382^{T} (AB429373)
Lactobacillus buchneri YIT 0077^{T} (AB429368)
Lactobacillus parabuchneri YIT 0272^{T} (AB429372)
Lactobacillus sunkii YIT 11161^{T} (AB366385)
Lactobacillus otakiensis YIT 11163^{T} (AB366386)
Lactobacillus kefiri YIT 0222^{T} (AB429371)
Lactobacillus salivarius ATCC 11741^{T} (AF089108)
Lactobacillus perolens L532^{T} (Y19167)
Lactobacillus fermentum ATCC 14931^{T} (M58819)
Lactobacillus reuteri DSM 20016^{T} (L23507)
Lactobacillus gasseri YIT 0192^{T} (AB008209)
Lactobacillus acidophilus YIT 0070^{T} (AB008203)
Lactobacillus delbrueckii subsp. *delbrueckii* YIT 0080^{T} (AB008207)
Lactobacillus vitulinus JCM 8228^{T} (AB210825)
0.05

Fig. 1. Phylogenetic tree based on 16S rRNA gene sequences showing the relationship of *L. saniviri* sp. nov. YIT 12363^{T} and *L.s senioris* sp. nov. YIT 12364^{T} with strains of closely related species.

The tree was constructed by the neighbor-joining method on the basis of a comparison of approximately 1390 bp and *Lactobacillus vitulinus* JCM 8228^{T} was used as the outgroup. Bootstrap values (%) based on 1000 replications are given at nodes. Bar, 5% sequence divergence.

Part 2B

Subspeciation of *Bifidobacterium longum* by multilocus approaches and amplified fragment length polymorphism: Description of *B. longum* subsp. *suillum* subsp. nov., isolated from the feces of piglets

2B.1. Introduction

Currently, the names of 91 species and 15 subspecies of the genus *Bifidobacterium* have been validly published (J.P. Euzéby: List of Prokaryotic names with Standing in Nomenclature; https://www.bacterio.net/). Among these species, *B. longum* is considered one of the most important contributors to host health and representative strains are frequently used as a probiotics (Desbonnet, et al. 2010; Puccio, et al. 2007; Xiao, et al. 2003). *B. longum* was described by Reuter (1963) using a strain isolated from adult feces. Later the species was recognized as one of the most prevalent bifidobacterial species in the gastrointestinal tract of human adults (Biavati, et al. 1984). *Bifidobacterium infantis* was also proposed by Reuter (1963) using strains that were prevalent in the gastrointestinal tract of infants (Biavati, et al. 1986; Matsuki, et al. 1998) while *Bifidobacterium suis* was described by Matteuzzi, et al. (1971) for strains isolated from pig feces.

In surveys of DNA–DNA relatedness, *B. longum*, *B. infantis* and *B. suis* have been shown to have hybridization rates of about 70% and higher (Lauer and Kandler 1983; Scardovi, et al. 1979), while possessing more than 97% 16S rRNA gene sequence identity (Miyake, et al. 1998), suggesting interrelationships at the species level (Stackebrandt and Goebel 1994). In addition, these species showed high average nucleotide identity (ANI) values ranging from 95.5 to 96.6%, which were higher than 95% cut-off value recommended for species demarcation (Kim, et al. 2014; Lugli, et al. 2014). Sakata, et al. (2002) unified *B. longum*, *B. infantis* and *B. suis* into the single species, *B. longum*, on the basis of DNA–DNA hybridization values, ribotyping and randomly amplified polymorphic DNA (RAPD) typing. They established three biotypes – ‘longum’, ‘infantis’ and ‘suis’ – for strains belonging to the respective former species. However, a variety of other genotypic techniques, including transaldolase-specific PCR-denaturing gradient gel electrophoresis (Requena, et al. 2002), comparison of the *recA*, *tuf* and *ldh* gene sequences (Roy and Sirois 2000; Ventura and

Zink 2003) and repetitive element sequence-based PCR fingerprinting (Masco, et al. 2003; Ventura, et al. 2003), as well as the results of PAGE experiments on soluble proteins (Biavati, et al. 1982), clearly allowed to discriminate these three species. Mattarelli, et al. (2008) assessed the published results in order to resolve the taxonomic dispute. In agreement with the International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of *Bifidobacterium*, *Lactobacillus* and related organisms, they have proposed that these three species should be reclassified into three subspecies, *B. longum* subsp. *longum*, *B. longum* subsp. *infantis* and *B. longum* subsp. *suis*. Given the wide applicability of the species, there is a strong need for reliable identification of strains of *Bifidobacterium* at species or subspecies level, improving studies of their ecological distribution and biological features.

The amplified-fragment length polymorphism (AFLP) method has been used to discriminate strains so as to verify the mother to infant transmission of *B. longum* subsp. *longum* (Makino, et al. 2011). Here, we chose to use AFLP, regarded as a reliable nucleic acid fingerprinting method (Tindall, et al. 2010) was used to investigate the subspecies of *B. longum* because of the success of this technique in classifying subspecies of *L. delbrueckii* (Tanigawa and Watanabe 2011). In addition, the multilocus sequence analysis or typing (MLSA or MLST) methods have been widely used to discriminate bacterial strains (Makino, et al. 2011; Tanigawa and Watanabe 2011). MLSA is based on the use of a phylogenetic analysis of nucleotide sequences of housekeeping genes of strains belonging to closely related species, allowing to determine phylogenetic clustering patterns. Concatenation of genes was shown to be extremely useful for precise bacterial phylogenetic analysis (Teichmann and Mitchison 1999). Ventura, et al. (2006) used seven housekeeping-gene sequences – of *clpC*, *dnaB*, *dnaG*, *dnaJ1*, *purF*, *rpoC* and *xfp* – to analyze the phylogenetic relationships of *Bifidobacterium* species; they recommended the use of a phylogeny based upon concatenated sequences to improve the identification of members of the *Bifidobacterium* at species level. Delétoile, et al. (2010) also confirmed that phylogenetic analysis based on the concatenated sequences of seven housekeeping genes (*clpC*, *fusA*, *gyrB*, *ileS*, *purF*, *rplB* and *rpoB*) indicated a distinct separation of *B. longum* subsp. *infantis* from the cluster composed of subspecies' *longum* and *suis*. MLST is used to analyze intraspecific diversity by comparing allelic profiles (Maiden, et al. 1998). A novel subspeciation of *L. delbrueckii* was revealed by using MLST based on the analysis of seven housekeeping genes (*fusA*, *gyrB*, *hsp60*, *ileS*, *pyrG*, *recA* and *recG*) (Tanigawa and Watanabe 2011). By using MLST based on seven housekeeping genes (*clpC*, *dnaG*, *dnaJ*, *fusA*, *gyrB*, *purF* and *rpoB*), Makino, et al. (2011) revealed the transmission of intestinal *B. longum* subsp. *longum* strains from mother to infant.

In this study, the results obtained by using the MLST method corresponded well to the results of AFLP with respect to the identification of 207 *B. longum* subsp. *longum* strains. It is very likely that continued whole genome sequencing results will assist in better quantifying and clarifying taxonomic and functional differences between strains of *B. longum*. The aim of the present study was to: (i) further examine the subspecies composition of *B. longum* using the above listed methodologies and (ii) adapt the current taxonomic structure of the species if shown necessary.

2B.2. Materials and methods

2B.2.1. Bacterial strains and growth conditions

A total of 25 bacterial strains assigned as *B. longum* subsp. *longum, B. longum* subsp. *infantis* or *B. longum* subsp. *suis* were obtained from the Culture Collection of the Yakult Central Institute (YIT; Tokyo, Japan) and the Bologna University Scardovi Collection of *Bifidobacteria* (Su; BUSCoB Bologna, Italy) (Table 4). All bacterial strains were grown at 37°C for 24 h in GAM broth (Nissui, Tokyo, Japan) supplemented with 0.5% glucose.

2B.2.2. Phenotypic characterization

Morphological, cultural and biochemical testing according to standard techniques was performed at 37°C unless otherwise stated. Gram staining and catalase activity were determined by using cells grown on modified GAM agar at 37°C for 2 days. Motility was tested in modified GAM soft agar (0.15%). Carbohydrate fermentation tests were carried out by using the API50 CHL system (bioMérieux, Marcy-l'Étoile, France) according to the manufacturer's instructions with some modifications (Watanabe, et al. 2009). Fructose-6-phosphate phosphoketolase (F6PPK) activity was determined according to the method described by Orban and Patterson (2000) with hexadecyltrimethylammonium bromide (0.45 mg/ml) as the detergent for cell membrane disruption. For urease activity assay, bacterial cells of overnight cultures grown in GAM broth supplemented with 0.5% glucose were inoculated into 4 ml of GAM broth supplemented with 0.5% glucose, 0.1 ml of 6 M urea solution (Toyobo Enzymes, Osaka, Japan) and 0.001% phenol red. After cultivation at 37°C for 4 days, urease activity was determined by the color change in the medium from yellow (pH < 6.8) to red (pH > 8.2). Cell-wall peptidoglycan was prepared and hydrolyzed according to the method of Schleifer and Kandler (1972). Cell-wall amino acids were analyzed by using a method described previously (Chao, et al. 2012). Cellular fatty acid methyl esters were obtained from cells grown on GAM broth supplemented with 0.5% glucose at 37°C for 1 day by saponification, methylation and extraction using the method of Miller (1982) with minor modifications (Kuykendall, et al. 1988). Cell protein for analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/MS) was performed as described previously (Sato, et al. 2011) with a slight modification. Briefly, 8 ml culture (incubated overnight) was centrifuged at 14,000 × g for 5 min. The resultant cell pellet was washed twice in double-distilled water before the cells were resuspended in TMA-1 buffer (10 mM Tris-HCl, 30 mM NH_4Cl, 10 mM $MgCl_2$, 6 mM 2-mercaptoethanol) and then centrifuged down at 10,000 × g for 2 min. The pellet produced was suspended in the buffer and ground with zirconia silica beads (0.1 mm in diameter: Biospec Products, Bartlesville, OK, USA) by using a FastPrep 24 Instrument (MP Biomedicals, Irvine, CA, USA) at a speed of 4.0 m/s for 30 s to obtain cell lysates. Each resulting lysate was mixed with a solution of sinapinic acid as the matrix at a concentration of 10 mg/ml in 50% acetonitrile with 1% trifluoroacetic acid. The resulting mixture was centrifuged, again at 10,000 × g for 2 min, before 0.5 µl supernatant solution was applied to a FlexiMass-DS disposable MALDI target (Shimadzu, Kyoto, Japan), airdried and then overlaid with 1 µl matrix solution. MALDI mass spectra in the range of 2000–16,000 molecular masses (*m*/*z*) were recorded, in positive linear mode, by averaging 500 individual laser shots made in an Axima Performance MALDI-TOF/TOF spectrometer (Shimadzu) equipped with a pulsed N_2 laser (λ = 337.1 nm; pulse width = 3 ns; frequency = 50 Hz). The spectra were calibrated externally by using adrenocorticotropic hormone (ACTH, clip18-39; *m*/*z* 2466.72) and apomyoglobin (*m*/*z* 16,952.276). The masses of reference peaks were determined by averaging three measurements.

2B.2.3. Genomic DNA extraction

For 16S rRNA and housekeeping-gene sequencing, the genomic DNAs of the strains were extracted by using a method described previously (Watanabe, et al. 2008). For determination of DNA–DNA relatedness and DNA G + C content, chromosomal DNA was extracted according to the method of Marmur (1961). To assess its G + C content, DNA was enzymatically degraded into nucleosides (Mesbah, et al. 1989) which were then separated and quantified by HPLC. DNA–DNA reassociation analyses were performed according to the microdilution well technique, using photobiotin for DNA labelling (Ezaki, et al. 1989) and using a SpectraMax M2 (Molecular Devices, Sunnyvale, CA, USA) for fluorescence measurements. Reciprocal hybridization experiments were performed for each pair of strains at 50°C for 2 h in the presence of 50% formamide, using biotinylated DNA and unlabeled ssDNA, which was bound non-covalently to microplate wells. The highest and lowest values of eight replicate wells were excluded and the mean of the remaining values was calculated for each experiment.

2B.2.4. Amplified fragment length polymorphism analysis

An AFLP analysis protocol as described previously (Makino, et al. 2011) was used, with the following modification. Bacterial DNA samples were digested with *Msp*I (Takara Bio, Shiga, Japan) and *Mse*I (New England BioLabs, Ipswich, MA, USA). The composition of the restriction reaction mixtures (5 µl) were as follows: 1 × NE buffer (New England BioLabs), 1 × T buffer (Takara Bio), 1 mg/ml bovine serum albumin, 20 U of *Msp*I, 10 U of *Mse*I and 29 ng of template DNA. The restriction reaction mixtures were incubated at 37°C for 2 h. For the next step, the ligation reaction mixture composition was as follows: 1 × T4 DNA ligase buffer, 10 µM *Msp*I adapter (5'-CTCGTAGACTGCGTACA-3', 5'-CGTGTACGCAGTCTAC-3') (Makino, et al. 2011) and 10 µM *Mse*I adapter (5'-GACGATGAGTCCTGAGG-3',5'-TACTCAGGACTCAT-3') (Vos, et al. 1995), 35 U T4 ligase (Takara Bio) and 5 µl of digested DNA in a total volume of 10 µl. The ligation reaction mixtures were incubated at 20°C for 2 h. Following the addition of 90 µl TE buffer, the digested and ligated DNA was used as the template for the preselective PCR. For preselective PCR amplification, the preselective primers for *Msp*I (5'-GACTGCGTACACGGA-3') and for *Mse*I (5'-GATGAGTCCTGAGTAA-3') were used. Ten microliters of reaction regent contained 10 mM of Tris–HCl (pH 8.3), 50 mM of KCl, 1.5 mM of $MgCl_2$, 250 μM of each dNTP, 2.5 μM of *Msp*I preselective primer, 2.5 μM of *Mse*I preselective primer, 0.25 U of *Taq* polymerase (Takara Bio) and 1 μl of diluted digested DNA. The PCR amplification program consisted of initial heating at 94°C for 2 min; 28 cycles of 94°C for 20 s, 56°C for 30 s and 72°C for 1.5 min; and a final extension at 72°C for 5 min. Following the addition of 990 μl TE buffer, the digested and ligated DNA was used as the template for the selective PCR.

For selective PCR, the selective primers for *Msp*I (5'-GACTGCGTACACGGAA-3') and for *Mse*I (5'-GATGAGTCCTGAGTAAT-3') (selective bases at the 3'-end are underlined) were used. The 5'-end of the *Msp*I selective primer was labelled with 6-carboxy-fluorescein (FAM). Ten microliters of reaction reagent contained 10 mM of Tris-HCl (pH 8.3), 50 mM of KCl, 1.5 mM of $MgCl_2$, 100 μM of each dNTP, 0.025 μM of FAM-labelled *Msp*I primer, 3 μM of *Mse*I selective primer, 0.25 U of *Taq* polymerase (Takara Bio) and 1 μl of 100-fold diluted preselective PCR product. The PCR amplification program consisted of initial heating at 94°C for 2 min and 13 cycles at 94°C for 20 s, 65°C for 30 s, and 72°C for 60 s, with a decrease in the annealing temperature of 0.7°C /cycle, followed by 24 cycles at 94°C for 20 s, 56°C for 30 s, and 72°C for 60 s; and a final extension at 72°C for 2 min. One microliter of the 10-fold diluted selective PCR product with TE buffer was mixed with 9.75 μl of Hi-Di™ formamide (Applied Biosystems, Foster City, CA, USA), 0.25 μl of GeneScan™ 500 LIZ® size standards (Applied Biosystems) and denatured at 95°C for 2 min. The samples were then analyzed with an ABI 3130xl Genetic Analyzer (Life Technologies, Carlsbad, CA, USA). The fragment pattern was extracted as binary data with GeneMapper (Life Technologies) on the basis of an analysis range of 51–500 bp and a threshold fluorescence value of 50 arbitrary units. The binary data were then converted to characteristic data (e.g., 0 and 1 were converted to A and T, respectively), and a cluster analysis was conducted by using unweighted pair group method with arithmetic mean (UPGMA) (Sneath and Sokal 1973) algorism with MEGA6 soft-ware (Tamura, et al. 2013).

2B.2.5. Housekeeping-gene sequence analysis

The conditions used for PCR amplification of the partial sequences of housekeeping genes and subsequent DNA sequencing have been described previously (Tanigawa and Watanabe 2011). For DNA amplification of the genes encoding ATP-dependent Clp protease ATP-binding subunit (*clpC*), DNA primase (*dnaG*), chaperone protein dnaJ1 (*dnaJ1*), heat shock protein 60 (*hsp60*), amidophosphoribosyltransferase (*purF*), RNA polymerase beta subunit (*rpoC*) and putative xylulose 5-phosphate/fructose 6-phosphate phosphoketolase (*xfp*), the primers listed in Table 5 were used. The primers clpC-F and clpC-R and hsp60-F and hsp60-R were based on positions 991–1733 and 286–818 of the *clpC* and *hsp60* genes, respectively, from *B. longum* subsp. *longum* NCC 2705 (GenBank accession no. AE014295), and the primers dnaJ1-F and dnaJ1-R were based on positions 195–586 of the *dnaJ1* gene from *B. longum* subsp. *infantis* ATCC 15697 (CP001095). Each 25 μl of reaction mixture contained 10 mM of Tris-HCl (pH 8.3), 50 mM of KCl, 1.5 mM of $MgCl_2$, 200 μM of each dNTP, 0.4 μM of each primer, 0.5 U of *Taq* polymerase (Takara Bio) and 10 ng of template DNA. The PCR conditions consisted of an initial heating step at 94°C for 2 min; 30 cycles of 94°C for 20 s, 57°C (for *dnaJ1* gene), 55°C (for *purF* gene) or 51°C (for the remaining housekeeping genes) for 20 s and 72°C for 20 s; and a final extension step at 72°C for 3 min. The amplicons were purified with an Ampure® XP kit (Beckman Coulter, Brea, CA, USA) and subsequently sequenced with an ABI PRISM BigDye® Terminator v3.1 cycle sequencing kit (Life Technologies). Multiple alignment of the sequences of the 25 strains was performed with MEGA6 software. Approximately 650, 890, 360, 420,760, 1090 and 190 bp of the seven housekeeping genes *clpC*, *dnaG*, *dnaJ1*, *hsp60*, *purF*, *rpoC* and *xfp*, respectively, of the 25 strains were used to construct a phylogenetic tree by using the neighbor-joining method (Saitou and Nei 1987), with Kimura's two-parameter model as a substitution model (Kimura 1980). The statistical reliability of the tree was evaluated by bootstrap analysis of 1000 replicates (Felsenstein 1985).

2B.2.6. Multilocus sequence analysis

To conduct MLSA the seven housekeeping-gene sequences were concatenated, yielding approximately 4360 bp of sequences. Phylogenetic trees based on the single gene sequences were constructed with MEGA6 software. Neighbor-Net analysis (Bryant and Moulton 2004) based on the concatenated sequences from the seven genes was performed with the SplitsTree4 program (Huson and Bryant 2006). Sequence similarities were corrected by using the Jukes-Cantor correction (Jukes and Cantor 1969).

2B.2.7. Multilocus sequence typing analysis

For MLST, the seven housekeeping-gene sequences were imported into BioNumerics software version 6.6 (Applied Maths, Sint-Martens-Latem, Belgium) and were aligned as described previously (Tanigawa and Watanabe 2011). The sequences obtained for each locus were compared for all isolates, and an allele number was given to each distinct sequence variant. Each isolate was defined by an allele profile or sequence type (ST) derived from the combination of alleles at the seven loci analyzed (Table 6). The same ST was used for the strains when they shared the same allelic profiles. The STs were identified by arbitrary numbers assigned in order of description. The alleles were then extracted and used to construct a minimum spanning tree. A clustering analysis was also performed using the UPGMA algorithm on a similarity matrix calculated based on a categorical coefficient on the position of mutations in the seven concatenated housekeeping gene sequences.

2B.3. Results

2B.3.1. Amplified fragment length polymorphism of 25 *B. longum* strains

The AFLP method is useful for grouping bacterial strains according to their genetic relationships. This method was used to measure the extent of subspecies diversity, and we constructed a dendrogram based on the AFLP profiles of the strains examined (Fig. 2). The 25 *B. longum* strains were separated into three major clusters (A, B and C). Cluster A consisted of 11 strains of *B. longum* subsp. *longum*. Cluster B consisted of seven strains of *B. longum* subsp. *suis* and was further divided into two subclusters which coincided with the strains' ability to produce urease: Subcluster B-1 comprised five urease-positive strains (YIT 4082^T, YIT 4108, Su 868, Su 903 and Su 923) and Subcluster B-2 comprised two urease-negative strains (Su 851^T and Su 864). Cluster C consisted of seven strains of *B. longum* subsp. *infantis*.

2B.3.2. Multilocus sequence analysis

By using a split network tree based on the concatenated sequences (4360 bp) of the seven genes (*clpC*, *dnaG*, *dnaJ1*, *hsp60*, *purF*, *rpoC* and *xfp*), the 25 strains of *B. longum* were divided into three major clusters, Cluster A, Cluster B and Cluster C, which corresponded to the subspecies *B. longum* subsp. *longum*, *B. longum* subsp. *suis* and *B. longum* subsp. *infantis*, respectively (Fig. 3). The result of the split network analysis revealed that Cluster B consisted of seven strains of *B. longum* subsp. *suis* and was divided into two subclusters: Subcluster B-1, which comprised the five urease-positive strains (YIT 4082^T, YIT 4108, Su 868, Su 903 and Su 923) and Subcluster B-2, which comprised the two urease-negative strains (Su 851^T and Su 864).

2B.3.3. Multilocus sequence typing analysis of allelic diversity

MLST analysis was used to examine the genetic diversity of strains of *B. longum* and to discriminate subspecies. The sequences of the seven loci were determined for the 25 *B. longum* strains, and 20 different STs were obtained. Of these, 16 were assigned to single strains. Three STs (ST1, ST19 and ST20) were assigned to two strains. ST18 was assigned to three strains (Table 6). The consensus sequence templates ranged in length from 186 bp (*xfp*) to 1093 bp (*rpoC*). Polymorphic sites are shown in Supplementary Fig. S4.

2B.3.4. Multilocus sequence typing for subspeciation of *B. longum*

The number of alleles per locus ranged from seven (*xfp*) to 18 (*dnaG*) (Table 6). By combining the seven gene loci, 20 STs were distinguished. ST inter-relationships were studied using the minimum spanning tree algorithm (Fig. 4). The 25 strains of *B. longum* were categorized into eight groups: Group 1 (ST1, ST2, ST3, ST4, ST7, ST8, ST9, ST12, ST15 and ST16), Group 2 (ST18 and ST19), Group 3 (ST20), Group 4 (ST5), Group 5 (ST13), Group 6 (ST10, ST11 and ST17), Group 7 (ST6) and Group 8 (ST14). Strains with STs that were separated by six or more allelic differences grouped separately. The UPGMA cluster analysis of the similarity matrix obtained using the categorical coefficient on all mutations in the seven concatenated housekeeping sequences, revealed that the 25 strains could be separated into three major clusters: Cluster A consisting of 11 strains of *B. longum* subsp. *longum*; Cluster B consisting of seven strains of *B. longum* subsp. *suis*; Cluster C consisting of seven strains of *B. longum* subsp. *infantis*. Cluster A was divided into two subclusters: cluster A-1 and cluster A-2 included four strains and seven strains of *B. longum* subsp. *longum*, respectively; Cluster B was also divided into two subclusters: cluster B-1 included five urease-positive strains (YIT 4082T, YIT 4108, Su 868, Su 903 and Su 923); cluster B-2 included two urease-negative strains (Su 851^{T} and Su 864) (Fig. 5).

2B.3.5. Phenotypic characteristics of *B. longum*

F6PPK activity was observed in all of the 25 *B. longum* strains used. The alkaline production based on urease activity was observed in five urease-positive strains of *B. longum* subsp. *suis* (YIT 4082^{T}, YIT 4108, Su 868, Su 903 and Su 923), whereas no color change observed in two urease-negative strains (Su 851^{T} and Su 864) of *B. longum* subsp. *suis* (Supplementary Fig. S5). Two specific peaks at around 6228 and 7730 *m*/*z* were observed in the MALDI-TOF MS spectra that discriminated the two urease-negative strains of *B. longum* subsp. *suis* (Su 851^{T} and Su 864) from all other strains investigated. Meanwhile, specific peaks were observed at around 10970 *m*/*z* for the five urease-positive strains of *B. longum* subsp. *suis* (YIT 4082^{T}, YIT 4108, Su 868, Su 903 and Su 923) and the type strains of *B. longum* subsp. *infantis* (YIT 12734^{T}).Strains Su 851^{T} and Su 864 did not display these peaks (Supplementary Fig. S6). Peptidoglycan structures of the two strains Su 851^{T} and Su 864 and two *B. longum* subspecies' type strains of YIT 4082^{T} and YIT 12734^{T}, were identical (L-Orn (Lys)-L-Ser-L-Ala-L-Thr-L-Ala, with the presence of Ala, Glu, Ser, Thr and Leu/Orn + Lys in ratio of 2:1:1:0.7:0.7; data not shown). Cellular fatty acid compositions of the two urease-negative strains of *B. longum* subsp. *suis* (Su 851^{T} and Su 864) and the three type strains of the three *B. longum* subspecies (YIT 4082^{T}, YIT 4021^{T} and YIT 12734^{T}) were almost identical, with the presence of major components of $C_{16:0}$ and $C_{18:1}$ $\omega 9c$ (Table 7).

2B.4. Discussion

One of our aims was to verify the current subspeciation of *B. longum* (*B. longum* subsp. *longum*, *B. longum* subsp. *infantis* and *B. longum* subsp. *suis*) using genotypic methods on a representative subset of *B. longum* strains. The subspeciation of *B. longum* was proposed by the International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of *Bifidobacterium*, *Lactobacillus* and related organisms and was based on an accumulation of published results on phenotypic and genotypic features (Mattarelli, et al. 2008). Various methods of discriminating strains of *B. longum* into subspecies have been proposed by many researchers. It is well known that analysis of DNA–DNA hybridization rates and 16S rRNA gene sequences of *B. longum* strains does not allow the clear separation of taxonomic entities at subspecies level. Consequently, MLSA has been proposed as an alternative to 16S rRNA gene analysis and has been recommended for improving the species-level identification of *Bifidobacterium*. The concatenated sequences of seven housekeeping genes (*clpC*, *dnaB*, *dnaG*, *dnaJ1*, *purF*, *rpoC* and *xfp*) have been used to discriminate type strains of the '*B. longum* group', which consists of *Bifidobacterium breve* and the three biotypes of *B. longum* ('longum', 'infantis' and 'suis' for strains belonging to *B. longum* subsp. *longum*, *B. longum* subsp. *infantis* and *B. longum* subsp. *suis*, respectively) (Ventura, et al. 2006). Delétoile, et al. (2010) found that MLSA based on the concatenated sequences of seven housekeeping genes (*clpC*, *fusA*, *gyrB*, *ileS*, *purF*, *rplB* and *rpoB*) discriminated *B. longum* subsp. *infantis* strains from a cluster comprising strains of *B. longum* subsp. *longum* and *B. longum* subsp. *suis*. However, the strains of *B. longum* subsp. *longum* were not separated from those of *B. longum* subsp. *suis*.

AFLP and MLSA techniques target different aspects of the genome. AFLP targets large genome rearrangements and compares the lengths of fragments of chromosomal DNA, whereas MLSA is based on the phylogenetic analysis of protein-coding housekeeping-gene sequences of strains belonging to closely related species. MLST, in contrast, is used to analyze intra-species diversity by comparing the allelic profiles of housekeeping genes. MLST, based on seven housekeeping-gene sequences (*fusA*, *gyrB*, *hsp60*, *ileS*, *pyrG*, *recA* and *recG*), has been used successfully to newly classify the subspecies of *L. delbrueckii* (Tanigawa and Watanabe 2011).

We therefore used the AFLP, MLSA and MLST methods to analyze 25 strains of *B. longum*, consisting of seven *B. longum* subsp. *infantis* strains, 11 *B. longum* subsp. *longum* strains and seven *B. longum* subsp. *suis* strains (five urease-positive and two urease-negative).

We confirmed that subspecies identification of the 25 *B. longum* strains with AFLP results corresponded well to results obtained with MLSA, using the concatenated sequences of seven housekeeping genes (*clpC*, *dnaG*, *dnaJ1*, *hsp60*, *purF*, *rpoC* and *xfp*). By these methods, the 25 strains were divided into three major clusters (A, B and C) that included two subclusters (B-1 and B-2). We also confirmed that the three major clusters could be easily identified in accordance with the locations of the respective type strains as *B. longum* subsp. *longum, B. longum* subsp. *suis* and *B. longum* subsp. *infantis,* respectively, and that the separation of *B. longum* subsp. *suis* strains into two subclusters corresponded to the ability of the strains to produce urease (Figs. 2 and 3). By using the MLST method we categorized the 25 strains of *B. longum* into eight groups. The 11 strains of *B. longum* subsp. *longum* were bundled into the same clonal complex (Group 1), which corresponded to Cluster A in the AFLP and MLSA dendrograms. The five urease-positive strains and two urease-negative strains in the *B. longum* subsp. *suis* were bundled into Group 2 and Group 3, respectively, which corresponded to subcluster B-1 and subcluster B-2, respectively, in the AFLP and MLSA dendrograms. However, the seven strains of *B. longum* subsp. *infantis* were separated into five groups (designated 4-8): Group 6, which comprised three strains [YIT 12734^T (ST10), YIT 12735 (ST11) and YIT 4081 (ST17)], and four disparate groups, namely Group 4 [YIT 11889 (ST5)], Group 5 [YIT 4019 (ST13)], Group 7 [YIT 11945 (ST6)] and Group 8 [YIT 4020 (ST14)] (Fig. 4). Although genetic diversity was observed in the strains of *B. longum* subsp. *infantis*, the fundamental results of the MLST analysis corresponded to the AFLP and MLSA results.

We thus confirmed that the results of our AFLP, MLSA and MLST analyses were consistent and that the results supported the proposal by Matsuki, et al. (1998) to reclassify the three biotypes of *B. longum* as three subspecies: *B. longum* subsp. *longum*, *B. longum* subsp. *infantis* and *B. longum* subsp. *suis*. In addition, we found that the two urease-negative *B. longum* strains (Su 851^T and Su 864) formerly assigned to *B. longum* subsp. *suis* should be allocated to a novel subspecies of *B. longum* for which the name *B. longum* subsp. *suillum* subsp. nov. is proposed. The new subspecies can be discriminated from *B. longum* subsp. *suis* in terms of differences in housekeeping-gene sequences and urease activity. Owing to the limited number of *B. longum* subsp. *suillum* strains, further study needs to be conducted with expanded number of strains to confirm whether the absence of urease activity is commonly shared within this subspecies. Urease activity is generally considered an important characteristic in bifidobacteria. The strongest ureolytic strains belong mainly to the former species '*B. suis*', whereas only a few belong to *Bifidobacterium aesculapii*, *B. breve*, *Bifidobacterium magnum* and *Bifidobacterium subtile*. Less than l0% of *B. breve* and *B. longum* subsp. *longum* strains are ureolytic, i.e., "human" *Bifidobacterium* species and *Bifidobacterium bifidum* are only weakly ureolytic (Crociani and Matteuzzi 1982; Modesto, et al. 2014). On the basis of urease activity, previous studies have demonstrated that 27 strains of '*B. suis*' could be allocated to two groups, consisting of 20 urease-positive strains and seven urease-negative strains, although all of these strains formed a strictly homologous group that had high levels of DNA homology (92–103%) (Matteuzzi and Crociani 1973). Results from our study presented here allow the discrimination of these subgroups into two separate subspecies (Table 8). The separation of these taxonomic subgroups may also correspond to an important functional property. Urease is known to play an important role in human neonatal development. The urea content of human milk (approximately 15% of total nitrogen) makes up for a deficiency in protein concentration resulting not only from the infant's own nitrogen demands but also from the metabolic requirements of its resident intestinal microbiota. Bacterial urease activity therefore contributes actively to the host anabolic processes in consequence of liberating ammonia from urea (Donovan and Lonnerdal 1989). Several studies, based upon the complete genome sequences of *B. longum* strains, have highlighted urease activity as part of an adaptation to the human gastrointestinal tract (Hao, et al. 2011; Schell, et al. 2002; Sela, et al. 2008; Wei, et al. 2010). Comparative genomic hybridization analysis and urease assays revealed that the urease gene cluster and its activity are furthermore conserved in all *B. longum* subsp. *infantis* and absent in all *B. longum* subsp. *longum* strains. This result implies that the *B. longum* has at least two distinct subspecies; *B. longum* subsp. *infantis*, specialized to utilize milk carbon (especially oligosaccharides) and *B. longum* subsp. *longum*, specialized for plant-derived carbon metabolism (LoCascio, et al. 2010).

2B.5. Description of novel subspecies

B. longum subsp. *suillum* (su.il'lum; L. neut. adj. *suillum*, pertaining to swine). Cells grown in modified GAM broth are rods of various shapes (0.4–0.7 µm × 1.5–3 µm in size), with rounded or tapered ends, sometimes curved, swollen and branched. They are Gram positive-staining, non-motile, asporogenous, F6PPK-positive, catalase-, oxidase- and urease-negative and anaerobic. After anaerobic growth at 37°C for 2 days, colonies on modified GAM agar are 2–3 mm in diameter; they are convex, white, opaque, smooth and circular with entire edges. The temperature range for growth is 20–45°C. The optimum temperature for growth is 25–37°C. Grows at pH 5.0–7.0 with an optimum but no growth at pH 9.5. Acid is produced from L-arabinose, D-ribose, D-xylose, D-galactose, D-glucose, D-fructose, D-mannose, methyl α-D-glucopyranoside, D-maltose, D-lactose, D-melibiose, D-sucrose, D-raffinose, and D-turanose. Acid is not produced from glycerol, erythritol, D-arabinose, L-xylose, D-adonitol, methyl α-D-xylopyranoside, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl β-D-mannopyranoside, *N*-acetylglucosamine, amygdalin, arbutin, salicin, D-cellobiose, D-trehalose, inulin, D-melezitose, starch, glycogen, xylitol, gentiobiose, D-lyxose, D-tagatose, D- or L-fucose, D- or L-arabitol, gluconate, 2- or 5-ketogluconate. Aesculin is hydrolyzed. Peptidoglycan type is A3β (L-Orn-L-Ser-L-Ala-L-Thr-L-Ala), with the presence of Ala, Glu, Ser, Thr and Leu/Orn + Lys in ratio of 2:1:1:0.7:0.7. The DNA G+C content of the type strain is 61.8 mol%. The type strain, Su 851^T (=DSM 28597^T = JCM 19995^T), and the reference strain Su 864 were isolated from the feces of piglets.

References

1. Biavati B, Castagnoli P, Crociani F, Trovatelli LD. 1984. Species of the *Bifidobacterium* in the feces of infants. Microbiologica **7**:341–345.
2. Biavati B, Castagnoli P, Trovatelli LD. 1986. Species of the genus *Bifidobacterium* in the feces of human adults. Microbiologica **9**:39–45.
3. Biavati B, Scardovi V, Moore WEC. 1982. Electrophoretic patterns of proteins in the genus *Bifidobacterium* and proposal of four new species. Int J Syst Bacteriol **32**:358–373.
4. Bryant D, Moulton V. 2004. Neighbor-Net: an agglomerative method for the construction of

phylogenetic networks. Mol Biol Evol **21**:255–265.

5. Chao SH, Kudo Y, Tsai YC, Watanabe K. 2012. *Lactobacillus futsaii* sp. nov., isolated from fu-tsai and suan-tsai, traditional taiwanese fermented mustard products. Int J Syst Evol Microbiol **62**:489–494.
6. Crociani F, Matteuzzi D. 1982. Urease activity in the genus *Bifidobacterium*. Ann Microbiol (Paris) **133 A**:417–423.
7. Delétoile A, Passet V, Aires J, Chambaud I, Butel MJ, Smokvina T, et al. 2010. Species delineation and clonal diversity in four *Bifidobacterium* species as revealed by multilocus sequencing. Res Microbiol **161**:82–90.
8. Desbonnet L, Garrett L, Clarke G, Kiely B, Cryan JF, Dinan TG. 2010. Effects of the probiotic *Bifidobacterium infantis* in the maternal separation model of depression. Neuroscience **170**:1179–1188.
9. Donovan SM, Lonnerdal BO. 1989. Non-protein nitrogen and true protein in infant formulas. Acta Paediatr **78**:497–504.
10. Ezaki T, Hashimoto Y, Yabuuchi E. 1989. Fluorometric deoxyribonucleic acid-deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane filter hybridization in which radioisotopes are used to determine genetic relatedness among bacterial strains. Int J Syst Bacteriol **39**:224–229.
11. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution (N Y) **39**:783–791.
12. Hao Y, Huang D, Guo H, Xiao M, An H, Zhao L, et al. 2011. Complete genome sequence of *Bifidobacterium longum* subsp. *longum* BBMN68, a new strain from a healthy Chinese centenarian. J Bacteriol **193**:787–788.
13. Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol **23**:254–267.
14. Jukes TH, Cantor CR. 1969. Evolution of protein molecules. pp. 21–132. In: Munro HN (Ed.), Mammalian protein metabolism. Academic Press, New York.
15. Kim M, Oh HS, Park SC, Chun J. 2014. Towards a taxonomic coherence between average

nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol **64**:346–351.

16. Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol **16**:111–120.
17. Kuykendall LD, Roy MA, O'Neill JJ, Devine TE. 1988. Fatty acids, antibiotic resistance, and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum*. Int J Syst Bacteriol **38**:358–361.
18. Lauer E, Kandler O. 1983. DNA-DNA homology, murein types and enzyme patterns in the type strains of the genus *Bifidobacterium*. Syst Appl Microbiol **4**:42–64.
19. LoCascio RG, Desai P, Sela DA, Weimer B, Mills DA. 2010. Broad conservation of milk utilization genes in *Bifidobacterium longum* subsp. *infantis* as revealed by comparative genomic hybridization. Appl Environ Microbiol **76**:7373–7381.
20. Lugli GA, Milani C, Turroni F, Duranti S, Ferrario C, Viappiani A, et al. 2014. Investigation of the evolutionary development of the genus *Bifidobacterium* by comparative genomics. Appl Environ Microbiol **80**:6383–6394.
21. Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A **95**:3140–3145.
22. Makino H, Kushiro A, Ishikawa E, Muylaert D, Kubota H, Sakai T, et al. 2011. Transmission of intestinal *Bifidobacterium longum* subsp. *longum* strains from mother to infant, determined by multilocus sequencing typing and amplified fragment length polymorphism. Appl Environ Microbiol **77**:6788–6793.
23. Marmur J. 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J Mol Biol **3**:208-IN1.
24. Masco L, Huys G, Gevers D, Verbrugghen L, Swings J. 2003. Identification of *Bifidobacterium* species using rep-PCR fingerprinting. Syst Appl Microbiol **26**:557–563.
25. Matsuki T, Watanabe K, Tanaka R, Oyaizu H. 1998. Rapid identification of human intestinal bifidobacteria by 16S rRNA-targeted species- and group-specific primers. FEMS Microbiol

Lett **167**:113–121.

26. Mattarelli P, Bonaparte C, Pot B, Biavati B. 2008. Proposal to reclassify the three biotypes of *Bifidobacterium longum* as three subspecies: *Bifidobacterium longum* subsp. *longum* subsp. nov., *Bifidobacterium longum* subsp. *infantis* comb. nov. and *Bifidobacterium longum* subsp. *suis* comb. nov.. Int J Syst Evol Microbiol **58**:767–772.
27. Matteuzzi D, Crociani F. 1973. Urease production and DNA-homology in the species *Bifidobacterium suis*. Arch Mikrobiol **94**:93–95.
28. Matteuzzi D, Crociani F, Zani G, Trovatelli LD. 1971. *Bifidobacterium suis* n.sp.: a new species of the genus *Bifidobacterium* isolated from pig feces. Z Allg Mikrobiol **11**:387–395.
29. Mesbah M, Premachandran U, Whitman WB. 1989. Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. Int J Syst Bacteriol **39**:159–167.
30. Miller LT. 1982. Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids. J Clin Microbiol **16**:584–586.
31. Miyake T, Watanabe K, Watanabe T, Oyaizu H. 1998. Phylogenetic analysis of the genus *Bifidobacterium* and related genera based on 16S rDNA sequences. Microbiol Immunol **42**:661–667.
32. Modesto M, Michelini S, Stefanini I, Ferrara A, Tacconi S, Biavati B, et al. 2014. *Bifidobacterium aesculapii* sp. nov., from the faeces of the baby common marmoset (*Callithrix jacchus*). Int J Syst Evol Microbiol **64**:2819–2827.
33. Orban JI, Patterson JA. 2000. Modification of the phosphoketolase assay for rapid identification of bifidobacteria. J Microbiol Methods **40**:221–224.
34. Puccio G, Cajozzo C, Meli F, Rochat F, Grathwohl D, Steenhout P. 2007. Clinical evaluation of a new starter formula for infants containing live *Bifidobacterium longum* BL999 and prebiotics. Nutrition **23**:1–8.
35. Requena T, Burton J, Matsuki T, Munro K, Simon MA, Tanaka R, et al. 2002. Identification, detection, and enumeration of human *Bifidobacterium* species by PCR targeting the transaldolase gene. Appl Environ Microbiol **68**:2420–2427.

36. Reuter G. 1963. Vergleichende Untersuchunge über die Bifidus-flora im Saüglings-und Erwachsenenstuhl. Zentralbl Bakteriol **191**:486–507.

37. Roy D, Sirois S. 2000. Molecular differentiation of *Bifidobacterium* species with amplified ribosomal DNA restriction analysis and alignment of short regions of the ldh gene. FEMS Microbiol Lett **191**:17–24.

38. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4**:406–425.

39. Sakata S, Kitahara M, Sakamoto M, Hayashi H, Fukuyama M, Benno Y. 2002. Unification of *Bifidobacterium infantis* and *Bifidobacterium suis* as *Bifidobacterium longum*. Int J Syst Evol Microbiol **52**:1945–1951.

40. Sato H, Teramoto K, Ishii Y, Watanabe K, Benno Y. 2011. Ribosomal protein profiling by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for phylogenety-based subspecies resolution of *Bifidobacterium longum*. Syst Appl Microbiol **34**:76–80.

41. Scardovi V, Casalicchio F, Vincenzi N. 1979. Multiple electrophoretic forms of transaldolase and 6-phosphogluconic dehydrogenase and their relationships to the taxonomy and ecology of the bifidobacteria. Int J Syst Bacteriol **29**:312–327.

42. Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G, et al. 2002. The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. Proc Natl Acad Sci U S A **99**:14422–14427.

43. Schleifer KH, Kandler O. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev **36**:407–477.

44. Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, et al. 2008. The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. Proc Natl Acad Sci U S A **105**:18964–18969.

45. Sneath P, Sokal R. 1973. Numerical taxonomy. The principles and practice of numerical classification. W. H. Freeman & Com-pany, San Francisco, CA.

46. Stackebrandt E, Goebel B. 1994. Taxonomic note: a place for DNA-DNA reassociation and

16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Bacteriol **44**:846–849.

47. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol **30**:2725–2729.
48. Tanigawa K, Watanabe K. 2011. Multilocus sequence typing reveals a novel subspeciation of *Lactobacillus delbrueckii*. Microbiology **157**:727–738.
49. Teichmann SA, Mitchison G. 1999. Is there a phylogenetic signal in prokaryote proteins? J Mol Evol **49**:98–107.
50. Tindall BJ, Rosselló-Móra R, Busse H-J, Ludwig W, Kämpfer P. 2010. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol **60**:249–266.
51. Ventura M, Canchaya C, Del Casale A, Dellaglio F, Neviani E, Fitzgerald GF, et al. 2006. Analysis of bifidobacterial evolution using a multilocus approach. Int J Syst Evol Microbiol **56**:2783–2792.
52. Ventura M, Meylan V, Zink R. 2003. Identification and tracing of *Bifidobacterium* species by use of enterobacterial repetitive intergenic consensus sequences. Appl Environ Microbiol **69**:4296–4301.
53. Ventura M, Zink R. 2003. Comparative sequence analysis of the *tuf* and *recA* genes and restriction fragment length polymorphism of the internal transcribed spacer region sequences supply additional tools for discriminating *Bifidobacterium lactis* from *Bifidobacterium animalis*. Appl Environ Microbiol **69**:7517–7522.
54. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res **23**:4407–4414.
55. Watanabe K, Fujimoto J, Sasamoto M, Dugersuren J, Tumursuh T, Demberel S. 2008. Diversity of lactic acid bacteria and yeasts in Airag and Tarag, traditional fermented milk products of Mongolia. World J Microbiol Biotechnol **24**:1313–1325.
56. Watanabe K, Makino H, Sasamoto M, Kudo Y, Fujimoto J, Demberel S. 2009. *Bifidobacterium mongoliense* sp. nov., from airag, a traditional fermented mare's milk product from Mongolia.

Int J Syst Evol Microbiol **59**:1535–1540.

57. Wei YX, Zhang ZY, Liu C, Zhu YZ, Zhu YQ, Zheng H, et al. 2010. Complete genome sequence of *Bifidobacterium longum* JDM301. J Bacteriol **192**:4076-4077.

58. Xiao JZ, Kondo S, Takahashi N, Miyaji K, Oshida K, Hiramatsu A, et al. 2003. Effects of milk products fermented by *Bifidobacterium longum* on blood lipids in rats and healthy adult male volunteers. J Dairy Sci **86**:2452–2461.

Tables and Figures

Table 4. Bacterial strains used in this study.

Species/Subspecies	Strain	Additional strain information	Source	Notes
B. longum subsp. *longum*	YIT 4021^T	ATCC 15707^T	Intestine of human infant	
	YIT 4037	ATCC 15708	Intestine of human infant	
	YIT 10936		Feces of human infant	
	YIT 10937		Feces of human infant	
	YIT 10938		Feces of human infant	
	YIT 11061		Intestine of human adult	
	YIT 11074		Feces of adlut human	
	YIT 11976	ATCC 55813	Feces of human infant	
	YIT 11977	ATCC 55814	Feces of human infant	
	YIT 12147		Intestine of human adult	
	YIT 12736	DSM 20097	Feces of calf	
B. longum subsp. *infantis*	YIT 4019	ATCC 15702	Intestine of human infant	
	YIT 4020	ATCC 25962	Intestine of human infant	
	YIT 4081		Feces of human infant	
	YIT 11889		Feces of human infant	
	YIT 11945		Feces of human infant	
	YIT 12734^T	DSM 20088^T	Intestine of human infant	
	YIT 12735	DSM 20218	Intestine of human infant	
B. longum subsp. *suis*	YIT 4082^T	JCM 1269^T, Su 859	Feces of piglet	urease positive
	YIT 4108	JCM 7139, Su 901	Feces of piglet	urease positive
	Su 868		Feces of piglet	urease positive
	Su 903		Feces of piglet	urease positive
	Su 923		Feces of piglet	urease positive
B. longum subsp. *suillum*	Su 851T		Feces of piglet	urease negative
subsp. nov	Su 864		Feces of piglet	urease negative

Table 5. Genes and primers used for MLSA and MLST.

Gene	Enzyme function	Position	Primer	Sequence (5'–3')	Analysis size (bp)	Reference
clpC	ATP-dependent Clp protease ATP-binding subunit	991–1733	clpC-F	GAGTACCGCAAGTACATCGAG	743	This study
			clpC-R	TCATCGTCGAACAGGAAC		
dnaG	DNA primase	408–780	DnaG-uni	CTGTGCCCGTTCCACGAC	373	Ventura, et al. 2006
			DnaG-rev	CTCGATGCGCAGGTCGCA		
dnaJ1	chaperone protein dnaJ1	195–586	dnaJ1-F	AGAKCGCCARAAGTACGACG	392	This study
			dnaJ1-R	TGAACTTCTTGCCGTCCACG		
hsp60	heat shock protein 60 (hsp60) gene	286–818	hsp60-F	CAGTCCCTGGTTCACGAGGG	533	This study
			hsp60-R	GCGCAGGACTTGAAGGT		
purF	amidophosphoribosyltransferase	609–1059	PurF-uni	CATTCGAACTCCGACACCGA	451	Ventura, et al. 2006
			PurF-rev	GTGGGGTAGTCGCCGTTG		
rpoC	RNA polymerase beta subunit	1194–1884	RpoC-uni	GTGCACTCGGTCCACAG	691	Ventura, et al. 2006
			RpoC-rev	CATGCTCAACAACGAGAAG		
xfp	putative xylulose 5-phosphate /fructose 6-phosphate phosphoketolase (xfp) gene	195–586	Xfp-uni	CTTCGGNCCNGANGAGAC	392	Ventura, et al. 2006
			Xfp-rev	AGCCACTTNGCNTGCTG		

Table 6. Typing of 25 strains by MLST, MLSA and AFLP.

Species/Subspecies	Allele code							ST	MLST group (Fig. 3)	Cluster type		AFLP group (Fig. 1)
	clpC	*dnaG*	*dnaJ1*	*hsp60*	*purF*	*rpoC*	*xfp*			MLST (Fig. 4)	MLSA (Fig. 2)	
B. longum subsp. *longum*												
YIT 4021[T]	1	14	1	1	12	12	1	15	1	A	A	A
YIT 4037	1	15	1	1	13	1	1	16	1	A	A	A
YIT 10936	1	1	1	1	1	1	1	1	1	A	A	A
YIT 10937	2	2	1	1	2	2	1	2	1	A	A	A
YIT 10938	1	1	1	1	1	1	1	1	1	A	A	A
YIT 11061	2	3	1	1	3	1	2	3	1	A	A	A
YIT 11074	2	3	1	1	2	3	3	4	1	A	A	A
YIT 11976	2	6	1	4	1	6	1	7	1	A	A	A
YIT 11977	1	7	1	4	2	7	1	8	1	A	A	A
YIT 12736	7	11	7	6	9	1	1	9	1	A	A	A
YIT 12147	1	8	4	1	6	1	1	12	1	A	A	A
B. longum subsp. *infantis*												
YIT 4019	8	12	8	7	10	10	4	13	6	C	C	C
YIT 4020	9	13	9	8	11	11	6	14	5	C	C	C
YIT 4081	6	16	3	8	14	13	4	17	6	C	C	C
YIT 11889	3	4	2	2	4	4	4	5	6	C	C	C
YIT 11945	4	5	3	3	5	5	5	6	6	C	C	C
YIT 12734[T]	5	9	5	5	7	8	4	10	4	C	C	C
YIT 12735	6	10	6	5	8	9	4	11	7	C	C	C
B. longum subsp. *suis*												
YIT 4082[T]	10	17	10	9	15	14	7	18	3	B-1	B-1	B-1
YIT 4108	10	17	10	9	16	14	7	19	3	B-1	B-1	B-1
Su 868	10	17	10	9	16	14	7	19	3	B-1	B-1	B-1
Su 903	10	17	10	9	15	14	7	18	3	B-1	B-1	B-1
Su 923	10	17	10	9	15	14	7	18	3	B-1	B-1	B-1
B. longum subsp. *suillum* subsp. nov.												
Su 851[T]	11	18	11	10	17	15	1	20	2	B-2	B-2	B-2
Su 864	11	18	11	10	17	15	1	20	2	B-2	B-2	B-2

Table 7. Cellular fatty acid composition of the novel strains and three type strains of *B. longum* subspecies.

Fatty acid	Su 851^T	SU 864	YIT 4082^T	YIT 12734^T	YIT 4021^T
Saturated					
$C_{12:0}$	–	–	–	1.06	–
$C_{14:0}$	7.70	5.70	9.14	9.83	2.84
$C_{14:0}$ DMA	5.34	2.38	8.34	4.02	3.08
$C_{16:0}$	24.46	24.40	28.57	14.26	24.68
$C_{16:0}$ DMA	0.82	0.47	0.66	0.26	1.78
$C_{18:0}$	1.92	3.58	2.42	1.38	4.20
$C_{18:0}$ DMA	0.44	0.31	0.20	–	1.00
Unsaturated					
$C_{18:1}$ ω 6*c*	6.13	–	–	–	2.09
$C_{18:1}$ ω 9*c*	5.25	12.51	11.27	24.62	11.82
$C_{18:1}$ ω 9*c* DMA	25.08	26.19	18.83	27.48	26.50
Cyclopropane					
$C_{19:0}$ cyclo 9,10	1.76	2.44	2.57	0.42	2.09
$C_{19:0}$ cyclo 9,10 DMA	8.36	6.70	5.00	1.02	5.51
Hydroxy fatty acid					
$C_{18:0}$ 12OH	1.50	1.37	0.59	1.17	1.67
Summed features[a]					
1	0.92	0.31	1.14	0.46	0.48
2	6.78	4.67	3.31	5.05	6.49
3	–	5.49	4.18	3.59	1.82
4	0.87	1.09	0.86	0.88	0.92
5	0.62	0.83	0.67	1.00	0.65

Strains: 1, *B. longum* subsp. *suillum* subsp. nov. Su 851^T; 2, *B. longum* subsp. *suillum* subsp. nov. Su 864; 3, *B. longum* subsp. *suis* YIT 4082^T; 4, *B. longum* subsp. *infantis* YIT 12734^T; 5, *B. longum* subsp. *longum* YIT 4021^T. Values are percentages of total fatty acids. The major components of cellular fatty acid are highlighted in bold.

[a] Summed features represent groups of two or more fatty acids that could not be separated by GLC with the MIDI system. Summed feature 1, $C_{13:1}$ at 12–13; $C_{14:0}$ aldehyde; $C_{11:1}$ 2OH. Summed feature 2, $C_{17:2}$ at 16.760; $C_{17:1}$ ω9*c*. Summed feature 3, $C_{18:1}$ ω7*c*; unknown 17.834. Summed feature 4, $C_{17:0}$ iso 3OH; $C_{18:2}$ DMA. Summed feature 5, unknown 18.622; $C_{19:0}$ iso.

–, not detected. DMA, dimethyl acetal.

Table 8. Major differential characteristics among the five piglet bifidobacterial strains and phylogenetically closely related bifidobacteria.

Characteristics	1	2	3	4	5	6	7	8
Acid production from:								
L-Arabinose	+	+	+	+	+	+	−	+
D-Ribose	+	+	−	−	−	−	+	+
D-Xylose	+	+	+	+	+	+	−	+
D-Mannose	+	+	+	+	+	+	+	−
N-Acethyl glucosamine	−	−	+	+	+	+	+	−
Amygdalin	−	−	−	−	−	w	−	−
Arbutin	−	−	−	w	−	−	+	−
Salicin	−	−	−	w	−	−	+	−
D-Treharose	−	−	−	−	−	w	w	+
Inulin	−	−	−	−	−	−	+	−
D-Melezitose	−	−	−	−	−	−	−	+
Gentiobiose	−	−	−	−	−	+	−	−
D-Turanose	+	+	+	+	+	+	+	−
L-Fucose	−	−	−	−	−	−	+	−
Aesculin hydrolysis	+	+	+	+	+	+	+	w
Growth at:								
20 °C	+	−	+	+	−	+	−	−
45 °C	+	−	+	+	+	+	−	−
pH 9.5	−	−	−	+	+	+	−	−
DNA G + C content (mol%)	61.8	61.9	ND	ND	ND	62.1	61.2	62.3

Strains: 1, *B. longum* subsp. *suillum* subsp. nov. Su 851^T; 2, *B. longum* subsp. *suillum* subsp. nov. Su 864; 3, *B. longum* subsp. *suis* Su 868; 4, *B. longum* subsp. *suis* Su 903; 5, *B. longum* subsp. *suis* Su 923; 6, *B. longum* subsp. *suis* YIT 4082^T; 7, *B. longum* subsp. *infantis* YIT 12734^T; 8, *B. longum* subsp. YIT 4021^T. Data were obtained in this study.

All eight strains produced acid from D-galactose, D-glucose, D-fructose, methyl α-D-glucopyranoside, D-maltose, D-lactose, D-melibiose, D-sucrose and D-raffinose. None of the strains produced acid from glycerol, erythritol, D-arabinose, L-xylose, D-adonitol, methyl β-D-xylopyranoside, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl α-D-mannopyranoside, D-cellobiose, starch, glycogen, xylitol, D-lyxose, D-tagatose, D-fucose, D- or L-arabitol, gluconate, 2- or 5-ketogluconate. +, positive. −, negative. w, weakly positive. ND, no data.

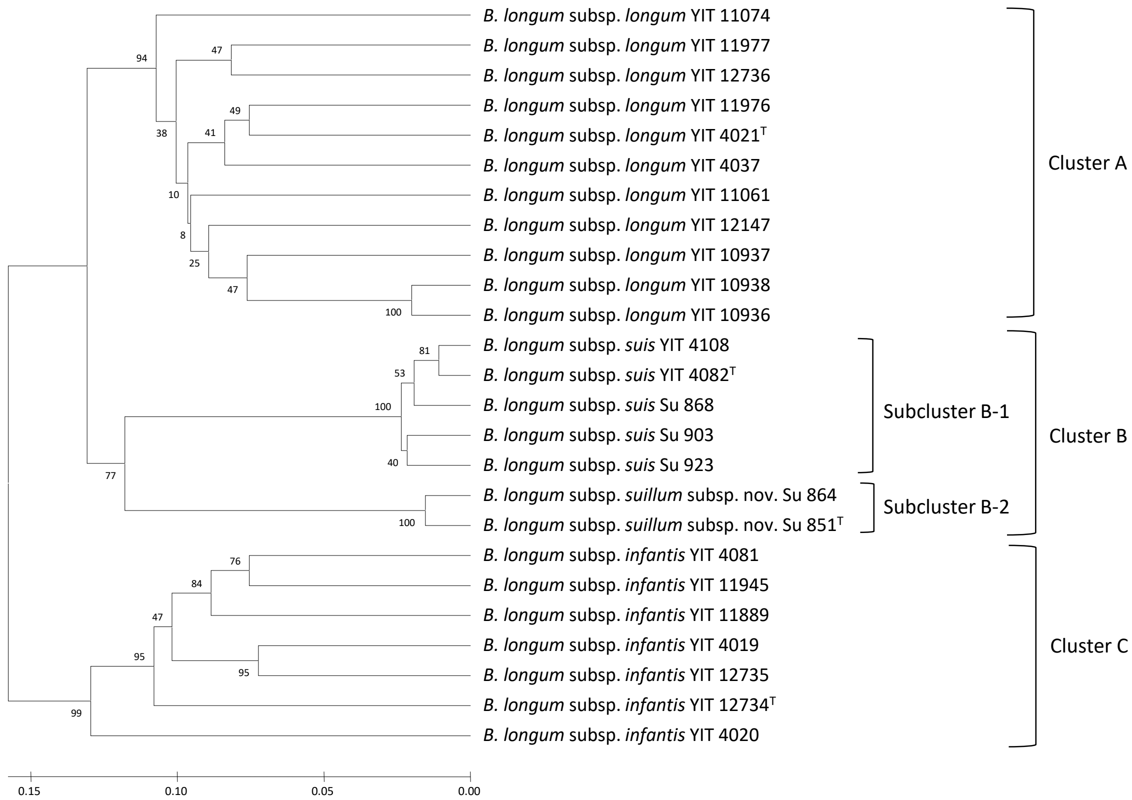


Fig. 2. Cluster analysis of digitalized AFLP fragment patterns of 25 strains of *B. longum*.

Dendrogram was constructed by using UPGMA. Bootstrap values based on 1000 replications are given at nodes. Bar, proportion of band mismatches.

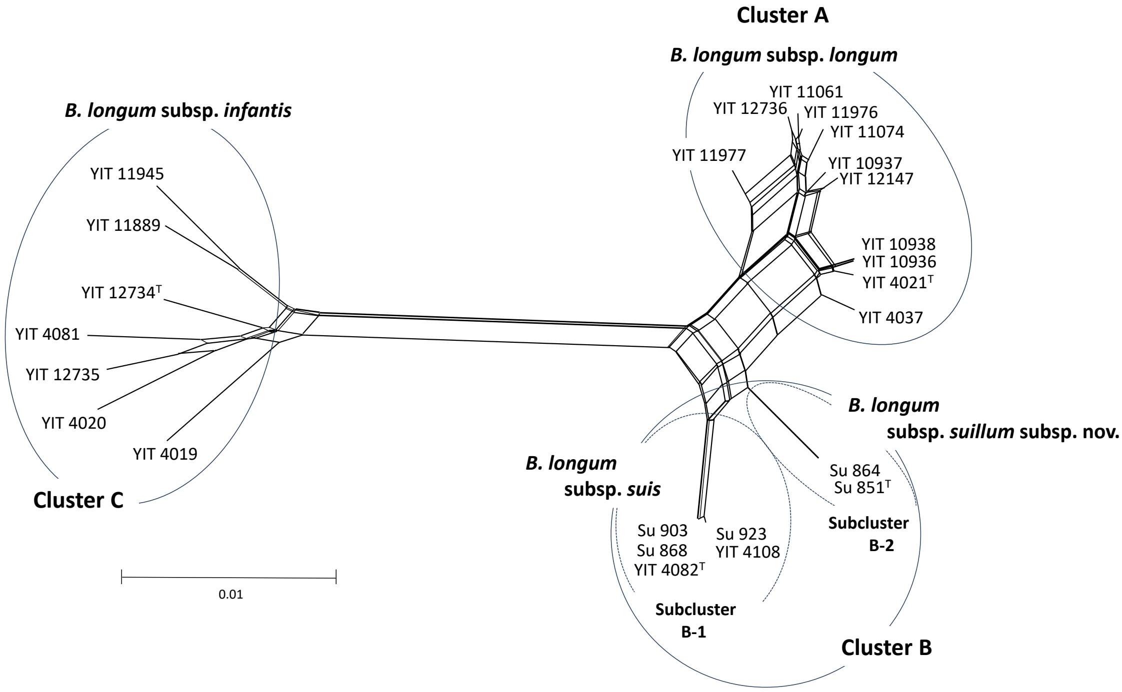


Fig. 3. Concatenated split network tree based on the seven genes.

The *clpC*, *dnaG*, *dnaJ1*, *hsp60*, *purF*, *rpoC* and *xfp* gene sequences (4360 bp) from 25 *B. longum* strains were concatenated and reconstructed using the SplitTree4 program. Sequence similarities were corrected using the Jukes-Cantor correction. Bar, 0.01 expected nucleotide substitutions per site.

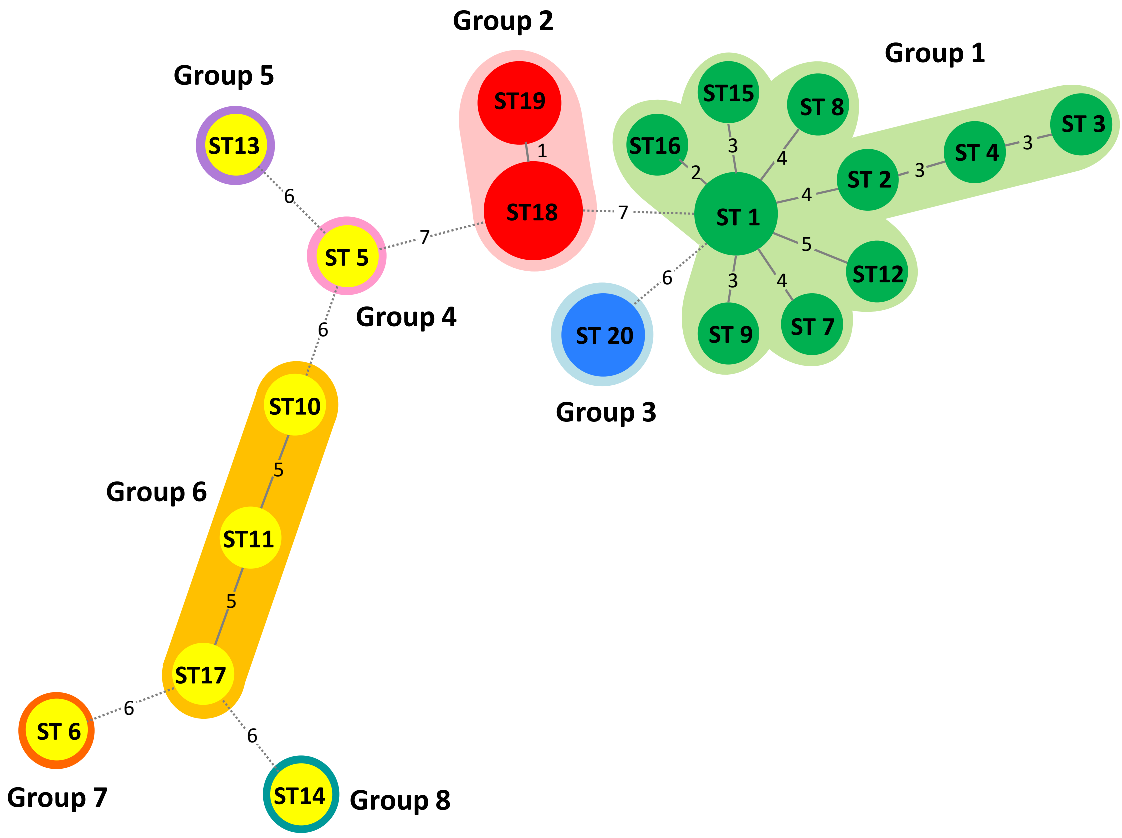


Fig. 4. Minimum spanning tree of 25 strains of B. longum based on sequence type (ST).

Each circle corresponds to the ST, and the circle size denoted the number of strains sharing the same ST. Numbers between the circles indicate the number of allelic differences between the profiles. When the distance between different STs was 6 or more, these STs were bundled as the same group.

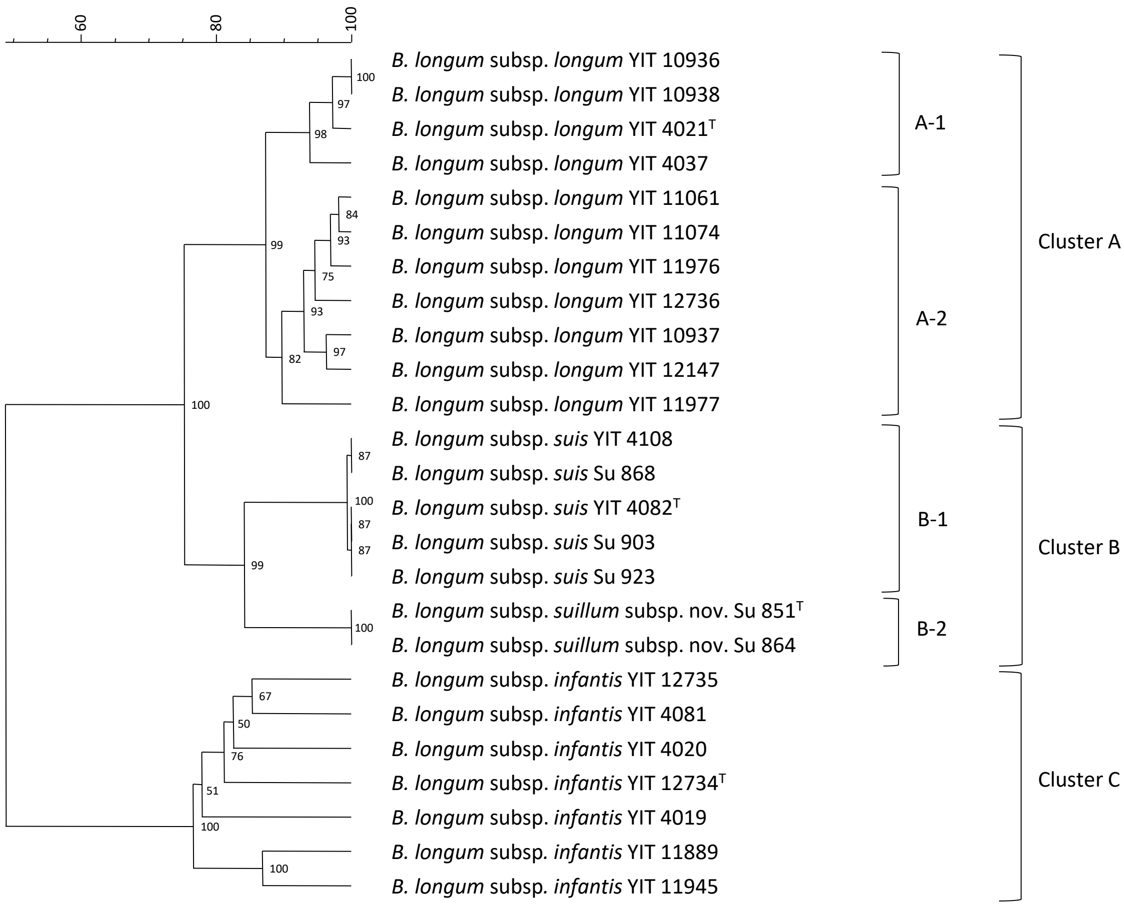


Fig. 5. Clustering analysis of MLST of 25 *B. longum* strains.

Dendrogram was constructed using UPGMA based on the concatenated discordant sequence of 7 housekeeping genes (320 bp). Bar, percentage sequence similarity.

Chapter 3: Study on a behavior of human gut *Bifidobacterium*

Part 3A

Long-term colonization exceeding six years from early infancy of *Bifidobacterium longum* subsp. *longum* in human gut

3A.1. Introduction

Several-hundred species of bacteria reside in the human gut (Guarner and Malagelada 2003) and a vast amount of evidence indicating the considerable influence of human gut microbiota on the health of the host has been accumulated (Clemente, et al. 2012). In addition, functional differences of bacterial strains regarding host health have also been reported, such as their virulence (Stubbs, et al. 2000), protective effects against pathogens (Fukuda, et al. 2011) and immunoregulatory properties (Khokhlova, et al. 2012).

Currently, the importance of the gut microbiota at the early stage of life has been well investigated. Moreover, it is suggested that the gut microbial composition during this period is associated with the risk of diseases (e.g. allergy, asthma and obesity) in the following life stages (Arrieta, et al. 2015; Sjögren, et al. 2009; Taveras, et al. 2009). The human gut microbiota develops just after delivery and the composition dynamically shifts throughout the lifetime of the host (Odamaki, et al. 2016). The composition of an infant's gut microbiota is influenced by various factors, such as the mode of delivery, diet, antibiotic usage during infancy and host genetics (Rodríguez, et al. 2015). In addition, Sharon *et al.* showed using a metagenomics approach that the shift of gut microbial composition occurs not only at the species level, but also at the strain level (Sharon, et al. 2013). They also showed that some bacterial strains belonging to *Staphylococcus epidermidis* and *Propionibacterium* spp. sustainably colonize in the infant gut from 15 up to 24 days of age. The sustainable existence of the same bacterial strain in early infancy for a certain period of time has also been confirmed for *Clostridium difficile* (Kubota, et al. 2016) and several species of the genus *Bifidobacterium* (Makino, et al. 2013; Milani, et al. 2015). However, it remains unclear whether a bacterial strain colonizing in the human gut during early infancy represents just a temporal resident for a limited period or persists to colonize the gut in the following life stage(s).

B. longum subsp. *longum* is a unique bifidobacterial species in the human gut detected at high prevalence and abundance, not only from infants, but also from adults and seniors (Gavini, et al. 2001; Ishikawa, et al. 2013; Makino, et al. 2013; Martin, et al. 2016). Several studies have shown that some strains of this species afford health-promoting potential to their host (Altmann, et al. 2016; Fukuda, et al. 2011; Takahashi, et al. 2006). Regarding their potential for continuous existence in the human gut, Shkoporov *et al.* reported the existence of two lineages of long-term colonizers that persist to colonize in the same subjects from 8 –16 months of age through 6 –10 years (Chaplin, et al. 2015; Shkoporov, et al. 2008). Considering the dynamic shift of gut bacterial strains in early infancy (Rodríguez, et al. 2015), there is little doubt that considerable selective stress exists during this period. Therefore, questions still remain regarding whether a strain belonging to *B. longum* subsp. *longum* colonizing in the human gut in early infancy may have the potential to constitute a long-term colonizer by overcoming the selective stresses.

To investigate this question, we conducted a follow-up study focusing on a Belgian cohort in which we have previously confirmed that a number of *B. longum* subsp. *longum* strains were transmitted from the mother's gut to that of the infant (Makino, et al. 2013, 2011) and were also shared between the infant's gut and the mother's postnatal breast milk (Makino, et al. 2015). In the present study, we confirmed the strain identity of *B. longum* subsp. *longum* isolates obtained from the fecal samples of an individual subject collected in both early infancy (in this study, the first six months of life) and childhood (approximately at six years of age). Furthermore, the analysis was expanded to the isolates obtained from their mothers' perinatal samples (prenatal fecal and postnatal breast milk samples) to investigate whether the mother-infant transmitted strains were able to become long-term colonizers.

3A.2. Material and methods

3A.2.1. Fecal sample collection

The current follow-up study was conducted in the area of Antwerp (Belgium) in 2016 (ISRCTN25216339). We recruited 49 subjects who had previously completed all the procedures of our previous study (ISRCTN66704989). The fecal samples used in this study were collected after at least five days of a washout period during which subjects were prohibited from taking fermented milk products. Following the methods of our previous studies (Makino, et al. 2013; Martin, et al. 2016), a portion of freshly voided feces was collected in a sterile glass tube containing 6 ml of anaerobic transport medium for the cultivation of *Bifidobacterium* and in a sterile tube containing 2 ml RNAlater® (Thermo Fisher Scientific, Waltham, MA, USA) for qPCR analysis. Samples were kept at 4°C after the collection and sent to the laboratory within one day after defecation. After arrival to the laboratory, the fecal sample for qPCR was washed twice with 1 ml phosphate buffered saline and stored at −20°C until DNA extraction.

3A.2.2. Bifidobacterial isolation and taxonomic identification

In the follow-up study, bifidobacteria were isolated from fecal samples and their DNAs were extracted, also according to Makino *et al.* (Makino, et al. 2013) with slight modifications. Briefly, serial dilutions of homogenized fecal samples were prepared with saline and inoculated onto a selective medium for *Bifidobacterium* (TOS propionate agar; Merck Co. Ltd., Darmstadt, Germany) supplemented with 50 μg/ml mupirocin (TOS-M agar), or TOS-M agar containing 16 μg/ml tetracycline (TOS-MT agar). Isolates obtained from TOS-MT agar were indicated with a “T” at the end of the strain ID. After anaerobic culturing at 37°C for 72 h, two to three colonies showing different colony morphologies were isolated. Additional single-colony isolation was carried out at least two times by using the same agar plate.

For DNA extraction, the purified bacterial isolates were anaerobically cultured in GAM medium (Nissui Co., Tokyo, Japan) supplemented with 1% glucose at 37°C for 24 h. Cellular DNA was extracted by means of phenol/glass bead extraction as previously described (Matsuki, et al. 2004) and used for subsequent taxonomic identifications. Initially, the species was determined based on partial nucleotide sequence of the 16S rRNA gene. The whole 16S rRNA gene was amplified by using the universal forward primer for *Bifidobacterium* BI8 and the universal reverse primer 15R as previously described (Miyake, et al. 1998). After purifying the amplicon using the Ampure® XP Kit (Beckman-Coulter, Brea, CA, USA), the nucleotide sequence of the target region was determined using the primers BI8 and 520R by using ABI PRISM BigDye® Terminator v3.1 chemistry (Life Technologies, Carlsbad, CA, USA) on a 3130xl Genetic Analyzer (Life Technologies). The determined sequence was searched against NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and the species was determined based on the highest score. If the bacterial isolate was identified as *B. longum*, the subspecies-specific PCR was conducted by using specific primers for *B. longum* subsp. *infantis* (BiINF-1 and BiINF-2) and those for *B. longum* subsp. *longum* (BiLON-1 and BiLON-2), as previously described (Matsuki, et al. 1999). The reaction mixture (25 µl) contained 10 mM Tris-HCl (pH 8.9), 50 mM KCl, 1.5 mM $MgCl_2$, 200 mM each dNTP, 0.5 U *Taq* DNA polymerase (Takara Bio, Shiga, Japan), 0.4 mM of each respective primer and 10 ng DNA template. The PCR amplification program was the same as that previously described (Matsuki, et al. 1999) with decreased PCR cycles (i.e. 30 cycles) and an extended extension step (i.e. 50 s). Specific primer sequences for PCR are listed in Table 9.

3A.2.3. Multilocus sequencing typing analysis

In order to remove duplicate isolates obtained from the same sample and to investigate the identity among the strains obtained from different samples, the bifidobacterial isolates were distinguished at strain level by multilocus sequence typing (MLST) analysis, as previously described (Makino, et al. 2011), based on the nucleotide sequences of seven housekeeping genes, i.e. *clpC* (class III stress response-related ATPase with chaperone activity), *dnaG* (DNA primase), *dnaJ* (chaperone protein DnaJ), *fusA* (GTP-binding protein chain elongation factor G), *gyrB* (the β subunit of DNA gyrase), *purF* (amidophosphoribosyltransferase) and *rpoB* (the β subunit of RNA polymerase).

In our previous studies (Makino, et al. 2013, 2011, 2015), bifidobacterial isolation was carried out from the subjects' fecal samples collected at seven time points of early infancy: meconium, 3, 7, 30 , 90 and 180 days of age, as well as after weaning (one week after the introduction of solids; at 140 ± 20 days for the subjects recruited in the follow-up study). In this study, we focused on the 12 subjects from whose fecal samples *B. longum* subsp. *longum* isolates were obtained in both early infancy and childhood (i.e., at the follow-up study conducted at approximately six years of age)

(Supplementary Table S1). We investigated the existence of long-term colonizers (i.e., strains persisting to exist in the same subject's gut from early infancy to childhood), as well as whether such strain(s) were transmitted from the mother's prenatal gut to the subject's gut or shared between the mother's postnatal breast milk and the subject's gut. Therefore, we confirmed the identities of the isolates obtained from the subjects' fecal samples collected in early infancy and childhood as well as from the mothers' perinatal samples (i.e., prenatal fecal sample collected twice with at least one-week interval before delivery and postnatal breast milk samples collected at 7 and 30 days after delivery) (Supplementary Tables S1 and S2).

The amplification of target genes for MLST analysis was conducted in 25 μl of reaction mixture containing 10 mM Tris-HCl (pH 8.9), 50 mM KCl, 1.5 mM $MgCl_2$, 200 mM each dNTP, 0.5 U *Taq* DNA polymerase (Takara Bio), 0.4 mM of each respective primer (Table 9) and 10 ng DNA template. The PCR amplification program consisted of an initial heating step at 94°C for 5 min; 30 cycles of 94°C for 30 s, 57°C for 30 s and 72°C for 1 min; and a final extension step at 72°C for 10 min. The procedures for amplicon purification and sequencing were the same as those for species identification except for the primers used in sequencing, which corresponded to those used for the amplification of the target genes. In addition to the sequences determined in this study and our previous study (Makino, et al. 2013), we extracted the corresponding housekeeping gene sequences from the genome sequences of the type strains of *B. longum* subsp. *infantis* (JCM 1222^T) and *B. longum* subsp. *longum* (JCM 1217^T) (Fukuda, et al. 2011) and used them in the following analysis.

The sequences were aligned for each gene based on the MUSCLE algorism v3.8.1 (Edgar 2004) mounted on GENETYX® Ver.12.0.5 (GENETYX, Tokyo, Japan). For each isolate, in total 2,902 bp (*clpC*, 479 bp; *dnaG*, 305 bp; *dnaJ*, 297 bp; *fusA*, 498 bp; *gyrB*, 396 bp; *purF*, 431 bp; *rpoB*, 496 bp) of nucleotide sequences were imported into BioNumerics® version 7.6 (Applied-Maths, Sint-Martens-Latem, Belgium). Each distinct gene sequence was assigned to an allelic number, and each unique combination of seven allelic numbers was assigned to an ST. The isolates showing the distinct ST for each sample at each sampling point were classified as the same strain.

3A.2.4. Clustering analysis based on allelic profiles

Clustering analysis was also carried out using BioNumerics® version 7.6. Together with the 96 representative strains (Supplementary Table S3), the type strains of *B. longum* subsp. *infantis* (JCM 1222^T) and *B. longum* subsp. *longum* (JCM 1217^T) were also included in this analysis. Among the 2,902 bp of the aligned nucleotide sequences of the seven housekeeping genes, 247 positions of the allelic profile were detected as those at which identical sequence was not shared among all of the strains used for this analysis (Supplementary Table S4). The categorical coefficient was calculated based on the allelic profiles and the dendrogram was constructed on the basis of the unweighted pair group method with arithmetic means (UPGMA) algorithm. The statistical reliability of the trees was evaluated by bootstrap analysis of 1,000 replicates (Felsenstein 1985) and the nodes replicated at more than 95% were regarded as statistically reliable.

3A.2.5. Amplified fragment length polymorphism analysis

Amplified fragment length polymorphism (AFLP) analysis was conducted following the methodology described previously (Makino, et al. 2011) with slight modifications. Seven representative strains belonging to ST 13-A, 30-A and 44-B (Supplementary Table S3) were subjected to AFLP analysis. Five microliter of restriction reaction mixtures contained 1 × CutSmart® buffer (New England BioLabs, Ipswich, MA), 5 U MseI (New England BioLabs), 5 U *Msp*I (New England BioLabs) and 30 ng of DNA template. The restriction reaction was conducted at 37°C for 2 h. Prior to ligation, equal amounts of adapters for *Mse*I and *Msp*I (Table 9) were separately mixed, denatured at 95°C for 5 min and left at room temperature for 5 min. Ligation was conducted in 10 µl of reaction mixtures containing 1 × T4 DNA ligase buffer (New England BioLabs), 2 µM *Mse*I adapters, 2 µM *Msp*I adapters, 40 U T4 DNA ligase (New England BioLabs) and 5 µl of digested DNA, with incubating at 20°C for 2h. The digested and ligated DNA was diluted 10-fold in Tris-EDTA buffer and used as template for the preselective PCR.

Preselective PCR was conducted in 10 µl of reaction mixture containing 10 mM Tris-HCl (pH 8.9), 50 mM KCl, 3 mM $MgCl_2$, 250 µM each dNTP, 0.25 U *Taq* DNA polymerase (Takara Bio), 2.5 µM of each preselective primer (Table 9) and 1 µl of template DNA. The PCR amplification program was the same as that previously described (Makino, et al. 2011). The amplicon was diluted 100-fold in Tris-EDTA buffer and used as template for the selective PCR.

Selective PCR was conducted in 10 µl of reaction mixture containing 10 mM Tris-HCl (pH 8.9), 50 mM KCl, 3 mM $MgCl_2$, 100 µM each dNTP, 0.25 U *Taq* DNA polymerase (Takara Bio), 30 nM of each selective primer (Table 9), and 1 µl of diluted amplicon of preselective PCR. The PCR amplification program was the same as that previously described (Makino, et al. 2011). Nine microliters of Hi-Di™ formamide (Life Technologies) and 1 µl of GeneScan™ 600 LIZ® size standards (Life Technologies) were mixed with 1 µl of the selective PCR products. Followed by denaturing at 95°C for 1 min, the selective PCR products were detected using a 3130xl Genetic Analyzer (Life Technologies). Output in FSA format was imported into BioNumerics® version 7.6 (Applied-Maths). AFLP analysis was conducted for the fragments ranging from 60 to 600 bp and a threshold of 1% was used for position tolerance. A UPGMA dendrogram was constructed based on Pearson correlation coefficients. The quality of each branch was evaluated by calculating the cophenetic correlation.

3A.2.6. Fecal DNA extraction and quantification of major fecal *Bifidobacterium*

The composition of the major fecal *Bifidobacterium* was analyzed by qPCR for the subjects' fecal samples collected in early infancy and childhood. According to our previous study (Martin, et al. 2016), fecal DNA was extracted from freeze-stored phosphate buffered saline-suspended fecal sample as described above and the abundances of the genus *Bifidobacterium,* along with that of seven species and three subspecies of fecal *Bifidobacterium* (*Bifidobacterium adolescentis*, *Bifidobacterium animalis* subsp. *lactis*, *B. bifidum*, *Bifidobacterium breve*, *Bifidobacterium catenulatum* group, *Bifidobacterium dentium*, *B. longum* subsp. *infantis* and *B. longum* subsp. *longum*) were measured using the primer sets listed in Table 9. PCR amplification and detection were performed with an ABI PRISM 7900HT Sequence Detection System and SDS software (version 2.4.1; Thermo Fisher Scientific, Waltham, MA, USA). 10 µl of the reaction mixture containing 10 mM Tris-HCl, pH 8.9; 50 mM KCl; 1.5 mM $MgCl_2$; 200 µM of each dNTP; 1:75,000 dilution of SYBR Green I (Thermo Fisher Scientific); 0.5 U of *Taq* DNA polymerase Hot start version (Takara Bio); 0.25 μM of each of the specific primers; and 1 μl of 10-fold, 100-fold, or 1,000-fold diluted template DNA. The amplification program consisted of one cycle at 94°C for 5 min; 40 cycles at 94°C for 20 s, 55°C for 20 s and 72°C for 50 s; and finally one cycle at 94°C for 15 s followed by the melting curve measurement using gradient heating increments of 0.2°C/s from 60°C to 95°C. Fluorescent products were detected during the last step of each cycle. For the subjects' fecal samples collected in early infancy, the abundance data measured in our previous study (Martin, et al. 2016) was used in this study.

3A.3. Results

3A.3.1 *B. longum* subsp. *longum* strains focused on this study

For 12 out of the 49 subjects recruited for this follow-up study, *B. longum* subsp. *longum* isolates were obtained from the fecal samples collected in both early infancy and childhood (Supplementary Table S1). For these 12 subjects, in total, 462 isolates were obtained (Supplementary Table S2), which came from the fecal samples collected in early infancy (243 isolates) and in childhood (46 isolates), as well as from their mothers' prenatal fecal (141 isolates) and postnatal breast milk samples (32 isolates). Based on the results from MLST analysis, all isolates were classified into 140 strains belonging to 73 sequence types (STs) as shown in Fig. 6. Strains with certain STs were distinctively detected from a single subject's fecal samples and the corresponding mother's perinatal samples. More than one strain was detected from some subjects' fecal samples collected in early infancy (max. five strains/subject at a time point) and childhood (max. four strains/subject at a time point), as well as some mothers' prenatal fecal samples (max. seven strains/subject at a time point) and the postnatal breast milk samples of the mother of subject 134 (two strains) (Table 10). Hereafter, strains originating from the same subjects' fecal sample and the corresponding mothers' perinatal samples sharing the same ST are defined as monophyletic strains.

3A.3.2 Comparison of sequence type

A total of 14 monophyletic strains (ST 12-A, 12-B, 16-A, 16-B, 16-C, 16-D, 18-A, 38-A, 38-B, 44-A, 130-A, 130-B, 130-C and 134-A) were detected from both the fecal samples of seven subjects in early infancy and the corresponding mothers' prenatal fecal samples (Fig. 6). In comparison, six monophyletic strains (ST 18-A, 44-B, 48-A, 129-A, 134-A and 134-B) were detected from both the fecal samples of five subjects in early infancy and the corresponding mothers' postnatal breast milk samples.

From the same subjects' fecal samples, three monophyletic strains (ST 13-A, 30-A and 44-B) were detected in both early infancy and childhood and were classified as long-term colonizers (Fig. 6). These long-term colonizers were detected from 120, 90 and 110 days of age, respectively. A monophyletic strain of a long-term colonizer (ST 44-B) was also isolated from the mother's postnatal breast milk sample collected at 30 days after delivery. No monophyletic strains obtained from a mother's prenatal fecal sample were detected from the subject's (offspring's) fecal sample in childhood, although 14 such strains were detected from seven subjects' fecal samples in early infancy.

3A.3.3. Comparison of allelic profiles

For the detailed comparison of the nucleotide sequences in the MLST loci, further clustering analysis was conducted. Among the strains sharing the same ST, we selected 96 representative strains detected at the earliest sampling points of the subjects' fecal samples collected in early infancy and childhood, as well as from their respective mothers' perinatal samples (Supplementary Table S3). The results from cluster analysis based on the 247 positions of allelic profiles (Supplementary Table S4) indicated that there were no distinct clusters composed of the long-term colonizers or the strains shared between subjects' fecal samples and their mothers' perinatal samples (Fig. 7). Although some clusters composed of the strains originating from a specific subject or sample type were suggested, the general composition of the dendrogram did not reflect the source subject or the isolated sample type of the strains.

Upon comparison of the allelic profiles between the strains detected from subjects' fecal samples in childhood and those of their closely related strain, 13-T2 (ST 13-I) and 16-T1 (ST 16-N) showed highly similar profiles (99.6% identity with one divergent position) compared to those of their relatives detected from the same subjects' fecal samples in early infancy, 1784 (ST 13-A) and

520 (ST 16-D), respectively (Fig. 7).

3A.4.4. Amplified fragment length polymorphism profiles of long-term colonizers

In order to investigate the genomic similarity among three monophyletic strains (ST 13-A, 30-A and 44-B) classified as long-term colonizers (Fig. 6), AFLP analysis was conducted. An average of 238 ± 14 fragments was detected from each of the seven representative strains. Based on the AFLP profile, the representative strains belonging to each ST were classified into distinct clusters with cophenetic correlation coefficients of > 99% (Fig. 8). In these clusters, highly similar AFLP profiles were observed among the strains (95.4%, 98.4% and 93.1-95.8% for ST 13-A, 30-A and 44-B, respectively), suggesting that considerable genomic similarity was conserved within each set of monophyletic strains.

3A.4.5. Composition of major fecal *Bifidobacterium*

In order to trace the population dynamics in the genus *Bifidobacterium,* including the long-term colonizers, the abundance of eight species and subspecies of the major fecal *Bifidobacterium* was measured by quantitative polymerase chain reaction (qPCR) for the subjects' fecal samples collected in early infancy and childhood (Supplementary Fig S7). The species and subspecies targeted in this analysis comprise most of the population of the genus *Bifidobacterium* at each time point, regardless of the subjects. Among the three subjects from whom monophyletic strains were detected in both early infancy and childhood (Subject IDs: 13, 30 and 44), *B. longum* subsp. *longum* was detected in early infancy from 90, 90 and 7 days of age, respectively, with the abundance of each strain being 8.08–9.65, 7.07–8.74 and 8.59–9.43 $\log_{10}$ cells/g feces, respectively (Fig. 9 and Supplementary Fig S7). Except for subject 13 at 180 days of age and subject 44 at 7 days of age, *B. longum* subsp. *longum* was not the most dominant component of *Bifidobacterium,* but rather co-existed with other predominant and dominant species (i.e., *B. adolescentis, B. bifidum, B. breve* and *B. catenulatum* group).

3A.4. Discussion

From 12 out of the 49 subjects, *B. longum* subsp. *longum* isolates were obtained from the fecal samples collected in both early infancy and childhood. Several strains were observed in some subjects' fecal samples collected in early infancy and childhood, as well as in the prenatal fecal samples of their mothers (Table 10 and Fig. 6). Although extensive intra-species bacterial diversity has been reported in the infant gut microbiota (Luo, et al. 2015; Sharon, et al. 2013), our result showed that the intra-species diversities of *B. longum* subsp. *longum* in the human gut exist not only in the infant but also in the child as well as in the prenatal mothers. Ellegaard and Engel have demonstrated some mechanisms hypothesized to facilitate strain co-existence (i.e., microniche differentiation, host selection, phage selection and cross-feeding/metabolic interdependency) (Ellegaard and Engel 2016). In addition, Odamaki *et al.* showed that *B. longum* subsp. *longum* strains are commonly transmitted among family members (Odamaki, et al. 2018). Together, these finding imply that several strains of *B. longum* subsp. *longum* that have been transmitted among the family members could coexist in the gut as the result of different selection pressures. Consistent with these findings, in the current study, strains with the same ST were not found to be shared among different subjects' fecal samples and their mothers' perinatal samples despite the intra-species diversities (Fig. 6).

Results from our current study revealed that three monophyletic strains (ST 13-A, 30-A and 44-B) persisted to colonize in the gut of individual subjects (Subject IDs: 13, 30 and 44) from early infancy for more than six years (Fig. 6). The close relation among the monophyletic strains was also supported by the results from the AFLP analysis (Fig. 9). Despite a high genomic similarity among each monophyletic stain, identical AFLP profiles were not observed among each of the strain combinations, suggesting the development of mutations during the colonization period in the human gut. This is consist with the previous report by Shkoporov *et al.* (Chaplin, et al. 2015). In addition to these long-term colonizers, the close relationship based on allelic profiles of two monophyletic strains detected from subjects 13 and 16 during childhood (ST 13-I and 16-N, respectively) implied that these strains evolved from common ancestors of the strains detected from the sample in early infancy of the same subjects (ST 13-A and 16-D, respectively) (Fig. 7).

In the present study, the long-term colonizers (ST 13-A, 30-A and 44-B) were not detected before 90 days of age, although 23 out of the 31 monophyletic strains isolated in early infancy were detected earlier than this age (Fig. 6). Additionally, in subject 44, the initial major component of *B. longum* subsp. *longum,* ST 44-A, was thought to be replaced by ST 44-B between 7 and 110 days of age. These results implied that earlier colonization by strains was not likely to contribute to the long-term colonization. However, it is difficult to explain long-term colonization based only upon colonization timing. For example, 14 monophyletic strains that were suggested to be vertically transmitted from mothers' gut were included in the "early" colonizers (Fig. 6). Considering that these strains were able to colonize in the mother's gut as a major component of *B. longum* subsp. *longum*, it had been supposed that these strains should also be capable of adapting to the subject's gut during childhood, which is more similar to the adult gut compared to that of the gut in early infancy.

In addition to the characteristics of bacterial strain (e.g. colonization timing, adhesion factor on the human gut epithelium and nutrient utilization), numerous other external factors should be considered that may affect the long-term colonization (e.g., host genetic background, immunological property, dietary habit and antibiotic usage, as well as competition with other gut microbiota). Recent studies suggested that among numerous factors, host diet has the greatest impact on microbial colonization in human gut (Donaldson, et al. 2016). A detailed analysis considering host diet and nutrient utilization properties for our long-term colonizers may provide information to explain the long-term colonization of this species in human gut. Owing to the limited the number of subjects and long-term colonizers in this study, we did not investigate the association of these factors with long-term colonization in order to avoid reaching a biased conclusion. Larger numbers of subjects and more detailed follow-up of subject backgrounds will be required to answer this question. Further analysis on the long-term colonizers, together with those reported in previous studies (e.g. genome comparison) (Chaplin, et al. 2015; Shkoporov, et al. 2008), would also likely contribute to our understanding of factors underlying the long-term colonization of this species in the human gut from the perspective of bacterial strain characteristics.

Several studies have reported an association between the gut microbial composition during early infancy and the risk of disease in the following life stages, including allergies, asthma and obesity (Arrieta, et al. 2015; Sjögren, et al. 2009; Taveras, et al. 2009). Our results suggested long-term colonization of bacterial strains from early infancy and implied that a bacterial strain colonized in the human gut during this period might have a longitudinal effect on the host health. Currently, some strains belonging to the genus *Bifidobacterium* (e.g. *B. animalis* subsp. *lactis*, *B. breve* and *B. longum* subsp. *infantis*) have been proposed for use as probiotics targeting infants (Di Gioia, et al. 2014). Our findings also showed the possibility that probiotics targeting infants might have longitudinal effects.

Although long-term colonizers were detected as the major component of *B. longum* subsp. *longum*, the source of these remained unclear. The cluster analysis based on the allelic profiles showed no phylogenetic relationship between long-term colonizers (Fig. 7), suggesting that the characteristics were not shared within a specific lineage of the strains belonging to this species. Although our study showed that one of the monophyletic strains (ST 44-B) was detected from breast milk (at 30 days after delivery) before the strain was detected from the subject fecal sample in early infancy (Fig. 6), we are unable to specify the precise origin of the strain. Possible origins of long-term colonizers would likely be family members of the subjects and the surrounding environment. For example, recent studies have shown a highly frequency of transmission of *B. longum* subsp. *longum* strains, not only from mother-to-infant, but also between other family members (Odamaki, et al. 2018), as well as the existence of a distinct microbiome in each family's home (Lax, et al. 2014).

The qPCR results showed that *B. longum* subsp. *longum* continued to be a dominant component of fecal *Bifidobacterium* (7.07 to 9.65 $\log_{10}$ cells/g feces) once this species colonized in the subject gut (Fig. 9). Since strain-specific quantification was not conducted in this study, we were unable to trace the actual dynamics for the abundance of long-term colonizers. Since only strains belonging to ST 30-A and 44-B were detected from the subjects' fecal samples after their colonization (Fig. 6), it was suggested that a considerable abundance of these strains persisted to colonize in the subject's gut for more than six years. In subject 44, the initial major component of *B. longum* subsp. *longum* (ST 44-A) was thought to be replaced by a long-term colonizer (ST 44-B) between 7 and 110 days of age. However, the precise timing of the replacement was unclear since no strain was detected at 30 days of age owing to limitations of the culturing approach. Moreover, for subject 13, because several strains were detected at 180 days and six years of age, the exact abundance of the long-term colonizer (ST 13-A) was not estimated at these time points.

Our qPCR results also revealed that the long-term colonizers were able to co-exist with other predominant species of *Bifidobacterium*. As previous studies have demonstrated the divergence of nutrient utilization (e.g. food-derived or human milk oligosaccharides and host-produced glycans) in *Bifidobacterium* at the species or strain level (De Vuyst, et al. 2014; Sela and Mills 2010; Turroni, et al. 2011), this implied that segregation of nutrient utilization may constitute one of the factors that might enable a long-term colonizer to persist among other predominant species of *Bifidobacterium*.

References

1. Altmann F, Kosma P, O'Callaghan A, Leahy S, Bottacini F, Molloy E, et al. 2016. Genome analysis and characterisation of the exopolysaccharide produced by *Bifidobacterium longum* subsp. *longum* 35624™. PLoS One **11**:e0162983.
2. Arrieta MC, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al. 2015. Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci Transl Med **7**:307ra152.
3. Chaplin AV, Efimov BA, Smeianov VV, Kafarskaia LI, Pikina AP, Shkoporov AN. 2015. Intraspecies genomic diversity and long-term persistence of *Bifidobacterium longum*. PLoS One **10**:e0135658.
4. Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. Cell **148**:1258–1270.
5. De Vuyst L, Moens F, Selak M, Rivière A, Leroy F. 2014. Summer Meeting 2013: growth and physiology of bifidobacteria. J Appl Microbiol **116**:477–491.

6. Di Gioia D, Aloisio I, Mazzola G, Biavati B. 2014. Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants. Appl Microbiol Biotechnol **98**:563–577.

7. Donaldson GP, Lee SM, Mazmanian SK. 2016. Gut biogeography of the bacterial microbiota. Nat Rev Microbiol **14**:20–32.

8. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5**:113.

9. Ellegaard KM, Engel P. 2016. Beyond 16S rRNA community profiling: intra-species diversity in the gut microbiota. Front Microbiol **7**:1475.

10. Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution (N Y) **39**:783–791.

11. Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, et al. 2011. Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature **469**:543–547.

12. Gavini F, Cayuela C, Antoine JM, Lecoq C, Lefebvre B, Membré JM, et al. 2001. Differences in the distribution of bifidobacterial and enterobacterial species in human faecal microflora of three different (children, adults, elderly) age groups. Microb Ecol Health Dis **13**:40–45.

13. Guarner F, Malagelada JR. 2003. Gut flora in health and disease. Lancet **361**:512–519.

14. Ishikawa E, Matsuki T, Kubota H, Makino H, Sakai T, Oishi K, et al. 2013. Ethnic diversity of gut microbiota: Species characterization of *Bacteroides fragilis* group and genus *Bifidobacterium* in healthy Belgian adults, and comparison with data from Japanese subjects. J Biosci Bioeng **116**:265–270.

15. Khokhlova EV, Smeianov VV, Efimov BA, Kafarskaia LI, Pavlova SI, Shkoporov AN. 2012. Anti-inflammatory properties of intestinal *Bifidobacterium* strains isolated from healthy infants. Microbiol Immunol **56**:27–39.

16. Kubota H, Makino H, Gawad A, Kushiro A, Ishikawa E, Sakai T, et al. 2016. Longitudinal investigation of carriage rates, counts, and genotypes of toxigenic *Clostridium difficile* in early infancy. Appl Environ Microbiol **82**:5806–5814.

17. Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, et al. 2014. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science **345**:1048–1052.

18. Luo C, Knight R, Siljander H, Knip M, Xavier RJ, Gevers D. 2015. ConStrains identifies microbial strains in metagenomic datasets. Nat Biotechnol **33**:1045–1052.

19. Makino H, Kushiro A, Ishikawa E, Kubota H, Gawad A, Sakai T, et al. 2013. Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota. PLoS One **8**:e78331.

20. Makino H, Kushiro A, Ishikawa E, Muylaert D, Kubota H, Sakai T, et al. 2011. Transmission of intestinal *Bifidobacterium longum* subsp. *longum* strains from mother to infant, determined by multilocus sequencing typing and amplified fragment length polymorphism. Appl Environ Microbiol **77**:6788–6793.

21. Makino H, Martin R, Ishikawa E, Gawad A, Kubota H, Sakai T, et al. 2015. Multilocus sequence typing of bifidobacterial strains from infant's faeces and human milk: are bifidobacteria being sustainably shared during breastfeeding? Benef Microbes **6**:563–572.

22. Martin R, Makino H, Cetinyurek Yavuz A, Ben-Amor K, Roelofs M, Ishikawa E, et al. 2016. Early-life events, including mode of delivery and type of feeding, siblings and gender, shape the developing gut microbiota. PLoS One **11**:e0158498.

23. Matsuki T, Watanabe K, Fujimoto J, Kado Y, Takada T, Matsumoto K, et al. 2004. Quantitative PCR with 16S rRNA-gene-targeted species-specific primers for analysis of human intestinal bifidobacteria. Appl Environ Microbiol **70**:167–173.

24. Matsuki T, Watanabe K, Tanaka R, Fukuda M, Oyaizu H. 1999. Distribution of bifidobacterial species in human intestinal microflora examined with 16S rRNA-gene-targeted species-specific primers. Appl Environ Microbiol **65**:4506–4512.

25. Milani C, Mancabelli L, Lugli GA, Duranti S, Turroni F, Ferrario C, et al. 2015. Exploring vertical transmission of bifidobacteria from mother to child. Appl Environ Microbiol **81**:7078–7087.

26. Miyake T, Watanabe K, Watanabe T, Oyaizu H. 1998. Phylogenetic analysis of the genus *Bifidobacterium* and related genera based on 16S rDNA sequences. Microbiol Immunol **42**:661–667.

27. Odamaki T, Bottacini F, Kato K, Mitsuyama E, Yoshida K, Horigome A, et al. 2018. Genomic diversity and distribution of *Bifidobacterium longum* subsp. *longum* across the human lifespan. Sci Rep **8**:85.

28. Odamaki T, Kato K, Sugahara H, Hashikura N, Takahashi S, Xiao JZ, et al. 2016. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. BMC Microbiol **16**:90.

29. Rodríguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N, et al. 2015. The composition of the gut microbiota throughout life, with an emphasis on early life. Microb Ecol Health Dis **26**:26050.

30. Sela DA, Mills DA. 2010. Nursing our microbiota: molecular linkages between bifidobacteria and milk oligosaccharides. Trends Microbiol **18**:298–307.

31. Sharon I, Morowitz MJ, Thomas BC, Costello EK, Relman DA, Banfield JF. 2013. Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. Genome Res **23**:111–120.

32. Shkoporov AN, Khokhlova E V, Kulagina E V, Smeianov V V, Kafarskaia LI, Efimov BA. 2008. Application of several molecular techniques to study numerically predominant *Bifidobacterium* spp. and *Bacteroidales* order strains in the feces of healthy children. Biosci Biotechnol Biochem **72**:742–748.

33. Sjögren YM, Jenmalm MC, Böttcher MF, Björkstén B, Sverremark-Ekström E. 2009. Altered early infant gut microbiota in children developing allergy up to 5 years of age. Clin Exp Allergy **39**:518–526.

34. Stubbs S, Rupnik M, Gibert M, Brazier J, Duerden B, Popoff M. 2000. Production of actin-specific ADP-ribosyltransferase (binary toxin) by strains of *Clostridium difficile*. FEMS Microbiol Lett **186**:307–312.

35. Takahashi N, Kitazawa H, Shimosato T, Iwabuchi N, Xiao J, Iwatsuki K, et al. 2006. An immunostimulatory DNA sequence from a probiotic strain of *Bifidobacterium longum* inhibits IgE production in vitro. FEMS Immunol Med Microbiol **46**:461–469.

36. Taveras EM, Rifas-Shiman SL, Belfort MB, Kleinman KP, Oken E, Gillman MW. 2009. Weight status in the first 6 months of life and obesity at 3 years of age. Pediatrics **123**:1177–1183.

37. Turroni F, Milani C, van Sinderen D, Ventura M. 2011. Genetic strategies for mucin metabolism in *Bifidobacterium bifidum* PRL2010: an example of possible human-microbe co-evolution. Gut Microbes **2**:183–189.

Table and Figures

Table 9. Primers used in this study.

Primer Name	Target	Sequence (5' - 3')	Reference
Amplification and sequencing of 16S rRNA gene for *Bifidobacterium*			
BI8	16S rRNA gene	GGGTTYCGATTCTGGCTCAGGATG	Miyake, et al. 1998
15R		AAGGAGGTGATCCARCCGCA	
520R		ACCGCGGCTGCTGGC	
Subspecies-specific PCR for *Bifidobacterium longum*			
BiINF-1	*Bifidobacterium longum* subsp. *infantis*	TTCCAGTTGATCGCATGGTC	Matsuki, et al. 1999
BiINF-2		GGAAACCCCATCTCTGGGAT	
BiLON-1	*Bifidobacterium longum* subsp. *longum*	TTCCAGTTGATCGCATGGTC	Matsuki, et al. 1999
BiLON-2		GGGAAGCCGTATCTCTACG	
MLST for *Bifidobacterium longum* subsp. *longum* strains			
Blon-clpC-F	*clpC*	CCTGAAGAAGGTGCTGAAGG	Makino, et al. 2011
Blon-clpC-R		TTCTCCTGCTTGTCGCGCAGT	
Blon-dnaG-F	*dnaG*	GTTGCCGTAGATTTGGGCTTGG	Makino, et al. 2011
Blon-dnaG-R		ATGACTTCGGTGTTCCGCAC	
Blon-dnaJ-F	*dnaJ*	GCTGAGCAAGAAGGAAGATCGC	Makino, et al. 2011
Blon-dnaJ-R		TGAACTTCTTGCCGTCCACGG	
Blon-fusA-F	*fusA*	CACCATCAAGGAGAAGCTGG	Makino, et al. 2011
Blon-fusA-R		ACGAGCTTGCCGTAGAACG	
Blon-gyrB-if1	*gyrB*	AAGTGCGCCGTCAGGGCTT	Makino, et al. 2011
Blon-gyrB-R		GTGTTCGCGAAGGTGTGCAC	
Blon-purF-F2	*purF*	CGGCTGAACTCGAAGAC	This study
Blon-purF-R2		GTTGAGCGCTTCCTTGAG	
Blon-rpoB-F	*rpoB*	AGACCGACAGCTTCGATTGG	Makino, et al. 2011
Blon-rpoB-R		AACACGATGGCGGACTGCTT	
AFLP analysis			
MseI adapter 1	Restriction site of MseI	TACTCAGGACTCAT	Makino, et al. 2011
MseI adapter 2		GACGATGAGTCCTGAG	
MspI adapter 1	Restriction site of MspI	CTCGTAGACTGCGTACA	Makino, et al. 2011
MspI adapter 2		CGTGTACGCAGTCTAC	
Preselective MseI	MseI adapter	GATGAGTCCTGAGTAA	Makino, et al. 2011
Preselective MspI	MspI adapter	GACTGCGTACACGGA	Makino, et al. 2011
Selective MseI-T	MseI adapter	GATGAGTCCTGAGTAAT	Makino, et al. 2011
Selective MspI-A	MspI adapter	FAM[a]-GACTGCGTACACGGAA	Makino, et al. 2011
Quantification of fecal *Bifidobacteirum*			
g-Bifid-F	Genus *Bifidobacterium*	CTCCTGGAAACGGGTGG	Martin, et al. 2016
g-Bifid-R		GGTGTTCTTCCCGATATCTACA	
BiADOg-1a	*Bifidobacterium adolescentis* group[b]	CTCCAGTTGGATGCATGTC	Martin, et al. 2016
BiADOg-1b		TCCAGTTGACCGCATGGT	
BiADO-2		CGAAGGCTTGCTCCCAGT	
Bflact2	*Bifidobacterium animalis* subsp. *lactis*	GTGGAGACACGGTTTCCC	Martin, et al. 2016
Bflact5		CACACCACACAATCCAATAC	
BiBIF-1	*Bifidobacterium bifidum*	CCACATGATCGCATGTGATTG	Martin, et al. 2016
BiBIF-2		CCGAAGGCTTGCTCCCAAA	
BiBRE-1	*Bifidobacterium breve*	CCGGATGCTCCATCACAC	Martin, et al. 2016
BiBRE-2		ACAAAGTGCCTTGCTCCCT	
BiCATg-1	*Bifidobacterium catenulatum* group[c]	CGGATGCTCCGACTCCT	Martin, et al. 2016
BiCATg-2		CGAAGGCTTGCTCCCGAT	
BiDEN-1	*Bifidobacterium dentium*	ATCCCGGGGGTTCGCCT	Martin, et al. 2016
BiDEN-2		GAAGGGCTTGCTCCCGA	
BiINF-1	*Bifidobacterium longum* subsp. *infantis*	TTCCAGTTGATCGCATGGTC	Martin, et al. 2016
BiINF-2		GGAAACCCCATCTCTGGGAT	
BiLON-1	*Bifidobacterium longum* subsp. *longum*	TTCCAGTTGATCGCATGGTC	Martin, et al. 2016
BiLON-2		GGGAAGCCGTATCTCTACG	

[a] 6-carboxyfluorescein; [b] The *B. adolescentis* group includes *B. adolescentis* genotypes A and B.

[c] The *B. catenulatum* group includes *B. catenulatum* and *Bifidobacterium pseudocatenulatum*.

Table 10. Count of monophyletic strains in the samples detected from 12 subjects and their mothers.

Subject ID	Detection of the representative strains[a]											
	Mother's perinatal samples				Subject's samples							
	Prenatal feces		Postnatal breast milk		Feces in early infancy							Feces in childhood
	1st Sample	2nd Sample	7 days	30 days	Meconium	3 days	7 days	30 days	90 days	After weaning[b]	180 days	≈ 6 years
4	2	2	–	–	–	–	–	–	–	1	–	3
12	7	5	–	–	–	1	2	2	–	ns	–	2
13	3	3	–	–	–	–	–	–	–	1	3	3
16	5	3	–	–	1	3	5	3	4	ns	1	2
18	1	1	1	–	–	1	1	1	1	–	–	1
30	2	2	–	–	–	ns	–	–	1	1	–	1
38	3	1	–	–	–	1	3	2	–	–	–	1
44	1	–	–	1	–	–	1	–	–	1	–	1
48	–	1	–	1	–	ns	–	2	–	–	–	2
129	–	–	1	1	1	1	1	1	–	–	–	1
130	4	3	–	–	–	3	–	–	–	–	–	4
134	1	–	–	2	–	2	2	2	–	–	–	2

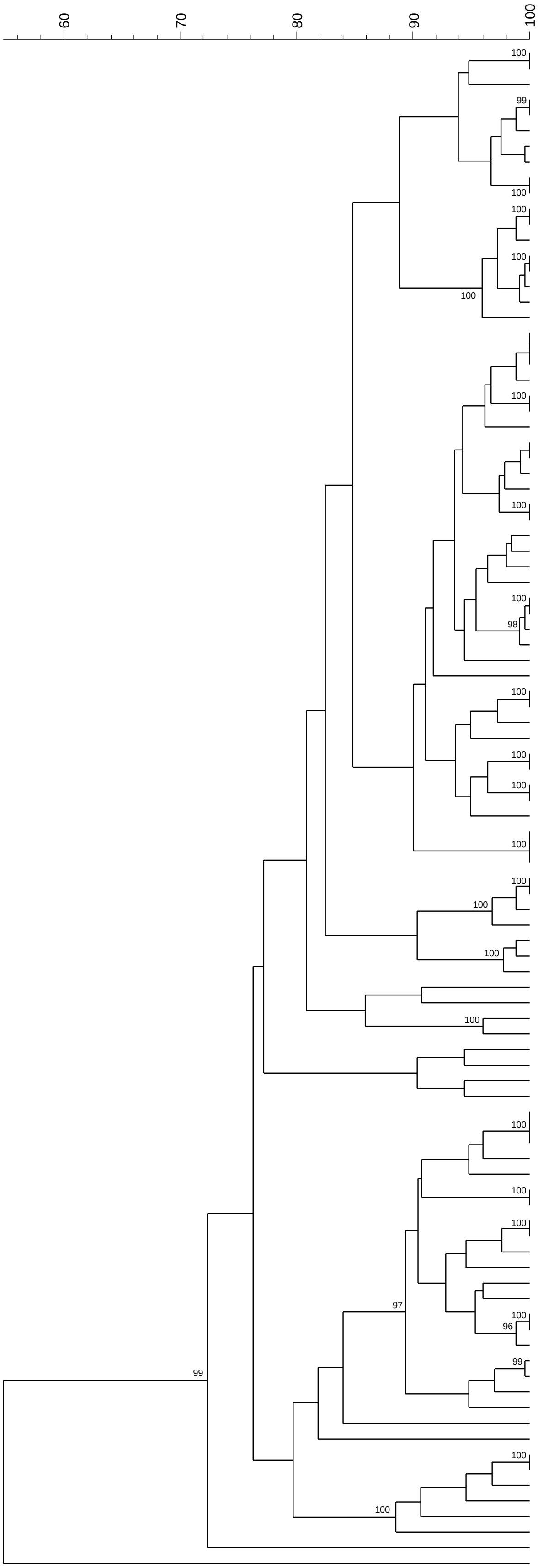
[a] ns, no sample was collected; −, no isolate was obtained.

[b] One week after the introduction of solids.

Subject ID	ST	Mother's perinatal samples				Subject's samples							
		Prenatal feces		Postnatal breast milk		Feces in early infancy							Feces in childhood
		1st Sample	2nd Sample	7 days	30 days	Meconium	3 days	7 days	30 days	90 days	After weaning[a]	180 days	≈ 6 years
4	4-A	○	○	–	–	–	–	–	–	–	–	–	–
	4-B	○	○	–	–	–	–	–	–	–	–	–	–
	4-C	–	–	–	–	–	–	–	–	–	○	–	–
	4-D	–	–	–	–	–	–	–	–	–	–	–	○
	4-E	–	–	–	–	–	–	–	–	–	–	–	○
	4-F	–	–	–	–	–	–	–	–	–	–	–	○
12	12-A	◐	–	–	–	–	◐	◐	◐	–		–	–
	12-B	◐	◐	–	–	–	–	◐	◐	–		–	–
	12-C	○	○	–	–	–	–	–	–	–		–	–
	12-D	○	–	–	–	–	–	–	–	–		–	–
	12-E	○	–	–	–	–	–	–	–	–		–	–
	12-F	○	–	–	–	–	–	–	–	–	ns	–	–
	12-G	○	–	–	–	–	–	–	–	–		–	–
	12-H	–	○	–	–	–	–	–	–	–		–	–
	12-I	–	○	–	–	–	–	–	–	–		–	–
	12-J	–	○	–	–	–	–	–	–	–		–	–
	12-K	–	–	–	–	–	–	–	–	–		–	○
	12-L	–	–	–	–	–	–	–	–	–		–	○
13	13-A	–	–	–	–	–	–	–	–	–	○	○	○
	13-B	○	○	–	–	–	–	–	–	–	–	–	–
	13-C	○	○	–	–	–	–	–	–	–	–	–	–
	13-D	○	–	–	–	–	–	–	–	–	–	–	–
	13-E	–	○	–	–	–	–	–	–	–	–	–	–
	13-F	–	–	–	–	–	–	–	–	–	–	○	–
	13-G	–	–	–	–	–	–	–	–	–	–	○	–
	13-H	–	–	–	–	–	–	–	–	–	–	–	○
	13-I	–	–	–	–	–	–	–	–	–	–	–	○
16	16-A	◐	◐	–	–	◐	◐	◐	–	–		–	–
	16-B	◐	–	–	–	–	◐	–	◐	◐		–	–
	16-C	◐	–	–	–	–	◐	–	–	–		–	–
	16-D	–	◐	–	–	–	–	◐	◐	–		–	–
	16-E	○	–	–	–	–	–	–	–	–		–	–
	16-F	○	○	–	–	–	–	–	–	–		–	–
	16-G	–	–	–	–	–	–	○	–	–	ns	–	–
	16-H	–	–	–	–	–	–	○	–	–		–	–
	16-I	–	–	–	–	–	–	○	–	–		–	–
	16-J	–	–	–	–	–	–	–	○	○		○	–
	16-K	–	–	–	–	–	–	–	–	○		–	–
	16-L	–	–	–	–	–	–	–	–	○		–	–
	16-M	–	–	–	–	–	–	–	–	–		–	○
	16-N	–	–	–	–	–	–	–	–	–		–	○
18	18-A	●	●	●	–	–	●	●	●	●	–	–	–
	18-B	–	–	–	–	–	–	–	–	–	–	–	○
30	30-A	–	–	–	–	–		–	–	○	○	–	○
	30-B	○	○	–	–	–	ns	–	–	–	–	–	–
	30-C	○	○	–	–	–		–	–	–	–	–	–
38	38-A	◐	◐	–	–	–	◐	◐	◐	–	–	–	–
	38-B	◐	–	–	–	–	–	◐	◐	–	–	–	–
	38-C	○	–	–	–	–	–	–	–	–	–	–	–
	38-D	–	–	–	–	–	–	○	–	–	–	–	–
	38-E	–	–	–	–	–	–	–	–	–	–	–	○
44	44-A	◐	–	–	–	–	–	◐	–	–	–	–	–
	44-B	–	–	–	◑	–	–	–	–	–	◑	–	◑
48	48-A	–	–	–	◑	–		–	◑	–	–	–	–
	48-B	–	○	–	–	–		–	–	–	–	–	–
	48-C	–	–	–	–	–	ns	–	○	–	–	–	–
	48-D	–	–	–	–	–		–	–	–	–	–	○
	48-E	–	–	–	–	–		–	–	–	–	–	○
129	129-A	–	–	◑	◑	◑	◑	◑	◑	–	–	–	–
	129-B	–	–	–	–	–	–	–	–	–	–	–	○
130	130-A	◐	◐	–	–	–	◐	–	–	–	–	–	–
	130-B	◐	◐	–	–	–	◐	–	–	–	–	–	–
	130-C	◐	–	–	–	–	◐	–	–	–	–	–	–
	130-D	○	–	–	–	–	–	–	–	–	–	–	–
	130-E	–	○	–	–	–	–	–	–	–	–	–	–
	130-F	–	–	–	–	–	–	–	–	–	–	–	○
	130-G	–	–	–	–	–	–	–	–	–	–	–	○
	130-H	–	–	–	–	–	–	–	–	–	–	–	○
	130-I	–	–	–	–	–	–	–	–	–	–	–	○
134	134-A	●	–	–	●	–	●	●	●	–	–	–	–
	134-B	–	–	–	◑	–	◑	◑	◑	–	–	–	–
	134-C	–	–	–	–	–	–	–	–	–	–	–	○
	134-D	–	–	–	–	–	–	–	–	–	–	–	○

Fig. 6. Detection points of the strains in the samples collected from 12 subjects and their mothers.

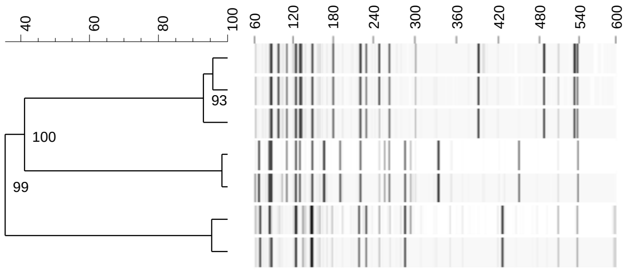
Circles indicate detection of the strains. Circle color represents the detection type of the monophyletic strain as follows: black, detected from both the subjects' fecal sample(s) in early infancy and their mothers' perinatal sample(s); half black (left), detected from both the subjects' fecal sample(s) in early infancy and their mothers' prenatal fecal sample(s); half black (right), detected from both the subjects' fecal sample(s) in early infancy and their mothers' postnatal breast milk sample(s); white, detected from only the subjects' fecal sample(s). A minus symbol means that no strain was detected. ns: no sample was collected. For each of the subject's samples, strains that were confirmed to be monophyletic strains are linked with a line. [a] One week after the introduction of solids.



Strain ID		ST	Subject ID	Isolated sample type[a]
2296		130-A	130	M
2317	*	130-A	130	S-I
724		16-J	16	S-I
281		16-A	16	M
401	*	16-A	16	S-I
336		12-I	12	M
1078		38-C	38	M
1687		38-D	38	S-I
2302		130-C	130	M
2315	*	130-C	130	S-I
297		16-C	16	M
441	*	16-C	16	S-I
298		16-E	16	M
326		16-D	16	M
520	*	16-D	16	S-I
16-T1		16-N	16	S-C
1314		16-L	16	S-I
521		16-H	16	S-I
1426		44-B	44	BM
1854	#†	44-B	44	S-I
44-1		44-B	44	S-C
4-4		4-E	4	S-C
1521		30-A	30	S-I
30-1	#	30-A	30	S-C
134-4		134-D	134	S-C
2297		130-B	130	M
2316	*	130-B	130	S-I
2295		130-D	130	M
2308		130-E	130	M
2220		129-A	129	BM
2285	†	129-A	129	S-I
144		13-D	13	M
4-6		4-F	4	S-C
B. longum subsp. *longum* (JCM1217^T)		LON	–	–
1800		4-C	4	S-I
1784		13-A	13	S-I
13-T1	#	13-A	13	S-C
13-T2		13-I	13	S-C
1847		13-F	13	S-I
31		4-A	4	M
12-4		12-K	12	S-C
286		16-B	16	M
435	*	16-B	16	S-I
16-1		16-M	16	S-C
1848		13-G	13	S-I
237		12-A	12	M
818	*	12-A	12	S-I
1082		38-A	38	M
1699	*	38-A	38	S-I
630		30-B	30	M
1120		18-A	18	S-I
1227	*†	18-A	18	BM
865		18-A	18	M
238		12-B	12	M
856	*	12-B	12	S-I
334		12-H	12	M
241		12-D	12	M
239		12-C	12	M
248		12-G	12	M
339		12-J	12	M
18-1		18-B	18	S-C
1362		48-B	48	M
244		12-E	12	M
246		12-F	12	M
527		16-I	16	S-I
1305		16-K	16	S-I
296		16-F	16	M
485		16-G	16	S-I
2452		134-A	134	BM
2476	*†	134-A	134	M
2488		134-A	134	S-I
631		30-C	30	M
141		13-C	13	M
2459		134-B	134	BM
2493	†	134-B	134	S-I
1688		38-B	38	S-I
1835	*	38-B	38	M
4-2		4-D	4	S-C
134-1		134-C	134	S-C
38-1		38-E	38	S-C
139		13-B	13	M
1268		44-A	44	S-I
977	*	44-A	44	M
129-4		129-B	129	S-C
130-5		130-H	130	S-C
130-6		130-I	130	S-C
130-1		130-F	130	S-C
130-3		130-G	130	S-C
35		4-B	4	M
220		13-E	13	M
1620		48-A	48	BM
1841	†	48-A	48	S-I
1844		48-C	48	S-I
48-1		48-D	48	S-C
13-1		13-H	13	S-C
48-2		48-E	48	S-C
12-5		12-L	12	S-C
B. longum subsp. *infantis* (JCM1222^T)		INF	–	–

Fig. 7. UPGMA dendrogram based on the allelic profiles.

The dendrogram was constructed based on 247 positions of allelic profiles of representative strains, as well as the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum*. *B. longum* subsp. *infantis* JCM 1222^{T} was used as the out-group. The scale bar shows the identical rate of allelic profile. Bootstrap values (%) based on 1,000 replicates are given for nodes replicated at more than 95%. [a] S-I, subject's fecal sample collected in early infancy; S-C, subject's fecal sample collected in childhood; M, mother's prenatal fecal sample; BM, postnatal breast milk sample. *Monophyletic strain pair detected from both subjects' fecal sample(s) in early infancy and their mothers' prenatal fecal sample(s). †Monophyletic strain pair detected from both subjects' fecal sample(s) in early infancy and their mothers' postnatal breast milk sample(s). #Monophyletic strain pair detected from subjects' fecal samples in both early infancy and childhood.



Strain ID	ST	Subject ID	Isolated sample type[a]	Detection time point[b]
1426	44-B	44	BM	30 days
1854	44-B	44	S-I	After weaning (110 days)
44-1	44-B	44	S-C	≈ 6 years
30-1	30-A	30	S-C	≈ 6 years
1521	30-A	30	S-I	90 days
13-T1	13-A	13	S-C	≈ 6 years
1784	13-A	13	S-I	After weaning (120 days)

Fig. 8. AFLP profiles of the representative strains belonging to three monophyletic strains classified as long-term colonizers.

The UPGMA dendrogram was constructed based on an average of 238 ± 14 fragments of the AFLP profile. Cophenetic correlation is given at each node. The scale bars for similarity score (%) and fragment length (bp) are shown above the dendrogram and AFLP profile image, respectively. [a] S-I, subject's fecal sample collected in early infancy; S-C, subject's fecal sample collected in childhood; BM, postnatal breast milk sample. [b] Calculated from the delivery date.

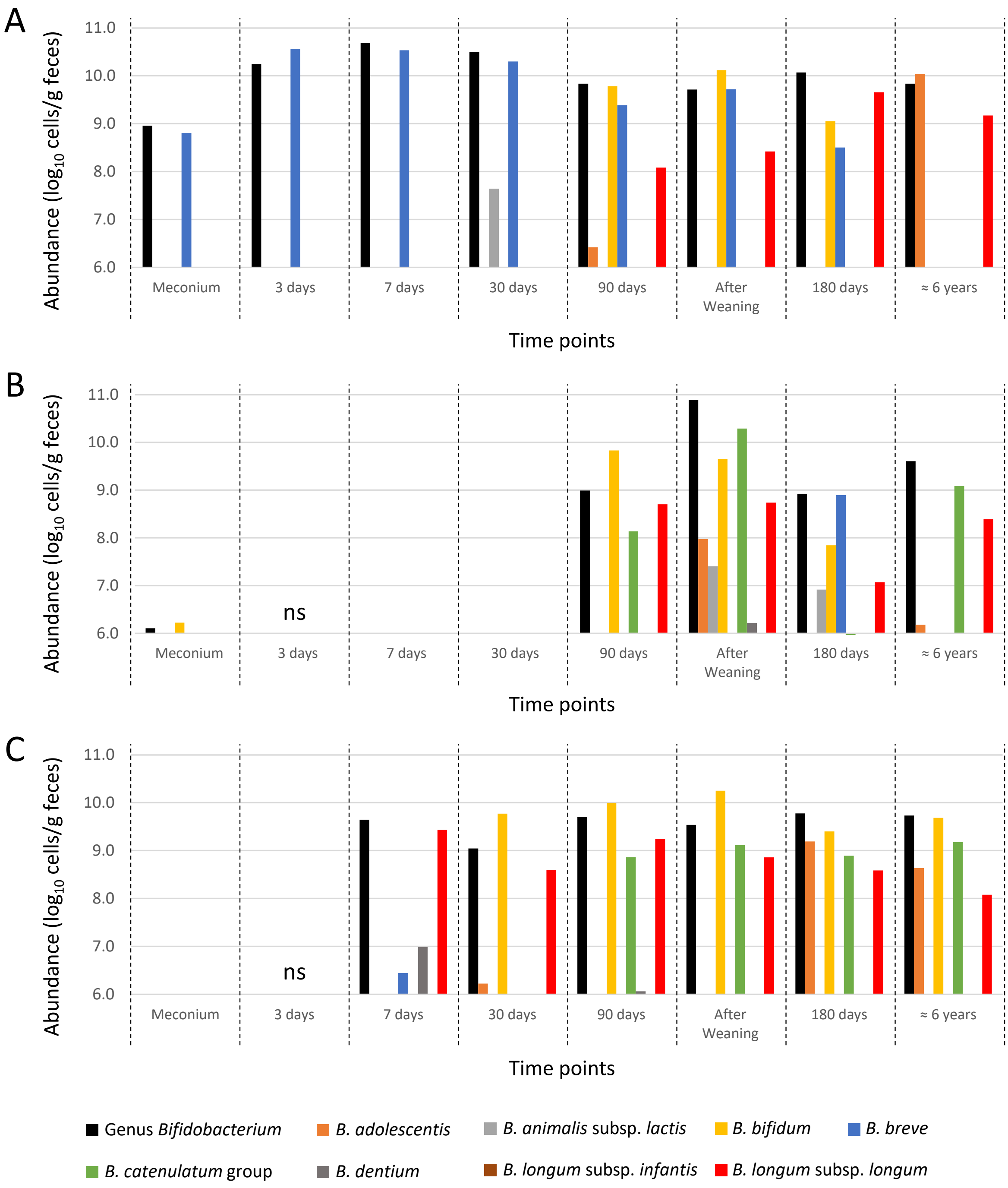


Fig. 9. Compositions of major fecal *Bifidobacterium* of the subjects colonized by long-term colonizers.

The fecal abundances of the genus *Bifidobacterium* along with seven species and three subspecies are shown for three subjects, (A) Subject 13, (B) Subject 30 and (C) Subject 44. ns: no sample was collected.

Chapter 4: Conclusion and comprehensive discussion

I conducted studies for the further understanding on one of the most complexed microbial ecosystem on Earth, human gut. Especially, I aimed at expanding the knowledge on human gut *Bifidobacterium* and *Lactobacillus*, from the point of views of their taxonomy and behavior.

In Chapter 2, I revised current taxonomy of *Lactobacillus* and *Bifidobacterium* with describing novel (sub)species and proposed a simplified subspeciation method for *Bifidobacterium longum*. Currently, 91 species and 15 subspecies as well as 253 species and 23 subspecies were described as “correct names” of genus *Bifidobacterium* and genus *Lactobacillus*, respectively (J.P. Euzéby: List of Prokaryotic names with Standing in Nomenclature; https://www.bacterio.net/). Since almost half of those species were described in this decade, the knowledge for their ecological feature have not sufficiently accumulated yet but will be revealed more in the future. As for the (sub)species which I described in this study, only limited number of strains is available, even so far. Therefore, the exploration of additional strains is still required to reveal the features shared at intra-(sub)species level, precisely.

As the further information for the features of *L. saniviri* and *L. senioris*, production of bacteriocin-like substance and inhibition of *Salmonella* growth were reported for a strain of *L. saniviri* isolated from fermented vegetable product (kimchi) prepared in USA (Kim, et al. 2015). In addition, Zúñiga, et al. (2020) reported fructan utilizing gene locus in the genome of *L. saniviri*, which is a limited feature in *Lactobacillus* strains. These results might show the anti-pathogenic potential and the unique niche based on fructan utilization of *L. saniviri* in human gut. Although the (sub)species which we described were not detected from human gut in the following studies, the existence of *L. saniviri* and *L. senioris* were reported by meta-16S rRNA gene analysis targeting on a fermented vegetable product (Yucha) prepared in a local community in China (Zhang, et al. 2016), suggesting their another ecological niche and the possibility that they utilize the plant derived nutrient. The draft genomes of the type strains of *L. saniviri* and *L. senioris* were determined along with those of other 211 species to encode genetic catalogue for modifying carbohydrates and proteins (Sun, et al. 2015). Although specific feature for these species have not been mentioned yet, the genomic information would contribute to understand the ecological function in the future study, e.g. providing new references for metagenomic approach.

Despite the absence of the further publication for *B. longum* subsp. *suillum*, number of researches have been conducted in this several years, focusing on human gut *B. longum* (mainly, *B. longum* subsp. *longum* and subsp. *infantis*) especially for their interaction to host's health (Amrane, et al. 2019; Olivares, et al. 2018; Zhou, et al. 2020). Correct subspeciation is the start point for a research using newly isolated strains or those in a culture collection without detailed identification. Although whole genome sequencing of bacterial strain is becoming easier and cheaper, recently, AFLP, MLSA and MLST methods developed in this study would continue to have the role as simplified methods for the subspeciation of this species. The methods developed in this study should contribute to the development of the research on *B. longum* with facilitating correct identification.

In Chapter 3, I showed long-term colonization of *B. longum* subsp. *longum* strains in the human gut from early stage of life and the possibility of their longitudinal effect on hosts' health. This perspective should be important for future studies on the bifidobacterial strains colonized in the human gut during this period and their effect on host's health. Complete explanation has not been provided for the long-term colonization factors for *B. longum* subsp. *longum* strain colonized in human gut during early infancy. A recent review reported that hosts' behavior, especially for dietary behavior, was the strongest factor for the succession of bacterial strain in infants (Quin and Gibson 2020). Moreover, dietary shift during early infancy was reported to correlate with the glycosyl hydrolase families and enzymes of *B. longum* strains (Kujawska, et al. 2020). Hence, the compatibility between the carbon metabolite property of a strain and host's dietary behavior would be important for the long--term colonization of this subspecies. On the other hand, it is still unclear whether there are endogenous long-term colonization factors of *B. longum* subsp. *longum* strains relating to the adhesion on human epithelial cells or to the protection from extracellular stresses (e.g. bile acid and antibiotics). The microbial molecules facilitating these properties were well studied for probiotic strain of *B. breve* UCC2003. Tad (tight-adherence) pili were reported as an essential for *in* *vivo* colonization of this strain on murine gut and the related gene cluster was widely conserved among *Bifidobacterium* (Motherway, et al. 2011). On the other hand, the protective function from bile and low pH and its influence on gut persistence was reported for the extracellular polysaccharide (EPS) of *B. breve* UCC2003 (Fanning, et al. 2012). Those compounds would be the candidates for the long-term colonization factor for *B. longum* subsp. *longum*. Further comparative genome analysis on long-term colonizers would help shed light on this question, e.g. with investigating the existence of the genetic loci relating to the production of those microbial compounds.

In summary, this study expanded the knowledge for the human gut microbiota, shedding light on the unknown components, providing the simplified methods for their detailed identification and showing the behavior during early stage of life, focusing on human gut *Bifidobacterium* and *Lactobacillus*. Recently, the knowledge have been accumulating for the interaction between human gut microbiota and environmental pollutants and the effect on host's health (Feng, et al. 2020). It has been shown that environmental pollutants affects the conformation and metabolic activity human gut microbiota, whereas human gut microbiota metabolites environmental pollutants and modulates its toxicity on host (Claus, et al. 2016). Since the effect of environmental pollutants on human health become global concern, the further understanding of human gut microbiota lying between them also would help elucidate the solution of this problem. One of the interesting approach is intestinal remediation with probiotics (El-Dalatony and Li 2020). In this review, several strains belonging to *Bifidobacterium* and *Lactobacillus* were reported to be able to diminish the accumulation of the pollutants and the toxicity via inhibiting the absorption of pollutants and enhancing gut barrier. Since endogenous *Bifidobacterium* and *Lactobacillus* also exist in human gut, it might be possible to elicit the remediation effect by modulating their composition in the human gut microbiota. In addition, adhesive activity to toxic metal (e.g. cadmium and lead) was reported for some strains of *Bifidobacterium* and *Lactobacillus* (Halttunen, et al. 2003, 2007). Therefore, further understanding on the mechanism on active cellular compounds also facilitating the remediation properties might leads to their application for the bioremediation.

Because human gut *Bifidobacterium* and *Lactobacillus* shall attract more researcher's interest in the future as beneficial components of human gut microbiota, the finding obtained in this study would also help the coming researches in this field.

References

1. Amrane S, Hocquart M, Afouda P, Kuete E, Pham TPT, Dione N, Ngom II, Valles C, Bachar D, Raoult D, Lagier JC. 2019. Metagenomic and culturomic analysis of gut microbiota dysbiosis during *Clostridium difficile* infection. Sci Rep **9**:1–8.
2. Claus SP, Guillou H, Ellero-Simatos S. 2016. The gut microbiota: a major player in the toxicity of environmental pollutants? Npj biofilms microbiomes.
3. El-Dalatony MM, Li X. 2020. Environmental pollutants that can be metabolized by the host (gut microbiota). pp. 145–168. In: Li X, Liu P (Eds.), Gut Remediation of Environmental Pollutant. Springer Singapore, Singapore.
4. Fanning S, Hall L, Cronin M, Zomer A, MacSharry J, Goulding D, Motherway M, Shanahan F, Nally K, Dougan G, van Sinderen D. 2012. Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. Proc Natl Acad Sci **109**:2108–2113.
5. Feng P, Xiao X, Zhou T, Li X. 2020. Effects of the bio-accumulative environmental pollutants on the gut microbiota. pp. 109–143. In: , Gut Remediation of Environmental Pollutants. Springer Singapore, Singapore.
6. Halttunen T, Kankaanpää P, Tahvonen R, Salminen S, Ouwehand AC. 2003. Cadmium removal by lactic acid bacteria. Biosci Microflora **22**:93–97.
7. Halttunen T, Salminen S, Tahvonen R. 2007. Rapid removal of lead and cadmium from water by specific lactic acid bacteria. Int J Food Microbiol **114**:30–35.
8. Kim JY, Young JA, Gunther NW, Lee J-L. 2015. Inhibition of *Salmonella* by bacteriocin-producing lactic acid bacteria derived from U.S. kimchi and broiler chicken. J Food Saf **35**:1–12.
9. Kujawska M, La Rosa SL, Pope PB, Hoyles L, McCartney AL, Hall LJ. 2020. Succession of *Bifidobacterium longum* strains in response to the changing early-life nutritional environment

reveals specific adaptations to distinct dietary substrates. bioRxiv.

10. Motherway MOC, Zomer A, Leahy SC, Reunanen J, Bottacini F, Claesson MJ, O'Brien F, Flynn K, Casey PG, Munoz JAM, Kearney B, Houston AM, O'Mahony C, Higgins DG, Shanahan F, Palva A, De Vos WM, Fitzgerald GF, Ventura M, O'Toole PW, Van Sinderen D. 2011. Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. Proc Natl Acad Sci U S A **108**:11217–11222.
11. Olivares M, Walker AW, Capilla A, Benítez-Páez A, Palau F, Parkhill J, Castillejo G, Sanz Y. 2018. Gut microbiota trajectory in early life may predict development of celiac disease. Microbiome **6**:1–11.
12. Quin C, Gibson DL. 2020. Human behavior, not race or geography, is the strongest predictor of microbial succession in the gut bacteriome of infants. Gut Microbes. Taylor and Francis Inc.
13. Sun Z, Harris HMB, McCann A, Guo C, Argimón S, Zhang W, Yang X, Jeffery IB, Cooney JC, Kagawa TF, Liu W, Song Y, Salvetti E, Wrobel A, Rasinkangas P, Parkhill J, Rea MC, O'Sullivan O, Ritari J, Douillard FP, Paul Ross R, Yang R, Briner AE, Felis GE, de Vos WM, Barrangou R, Klaenhammer TR, Caufield PW, Cui Y, Zhang H, O'Toole PW. 2015. Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. Nat Commun **6**:8322.
14. Zhang J, Wang X, Huo D, Li W, Hu Q, Xu C, Liu S, Li C. 2016. Metagenomic approach reveals microbial diversity and predictive microbial metabolic pathways in Yucha, a traditional Li fermented food. Sci Rep **6**:1–9.
15. Zhou C, Fang X, Xu J, Gao J, Zhang L, Zhao J, Meng Y, Zhou W, Han X, Bai Y, Li Z, Zou D. 2020. *Bifidobacterium longum* alleviates irritable bowel syndrome-related visceral hypersensitivity and microbiota dysbiosis via Paneth cell regulation. Gut Microbes 1–17.
16. Zúñiga M, Yebra MJ, Monedero V. 2020. Complex oligosaccharide utilization pathways in *Lactobacillus*. Curr Issues Mol Biol **40**:49–80.

Acknowledgements

Here I show my sincere appreciation to Dr. Koichi Watanabe and Dr. Junji Fujimoto, for their precious advices and discussions for my study. Without their supports and encouragements, this thesis was never written.

I would also like to express my sincere gratitude toward Prof. Masaaki Morikawa, Prof. Manabu Fukui, Assistant Prof. Kenji Washio and Assistant Prof. Hisaya Kojima for their many supports and discussions in the fulfillment of this thesis.

This study was carried out in Yakult Central Institute (Tokyo, Japan) and Yakult Honsha European Research Center for Microbiology VOF (Gent, Belgium), and partially received support from Danone Nutricia Research, Utrecht (Utrecht, The Netherlands). I appreciate deeply Dr. Fumiyasu Ishikawa (director of Yakult Central Institute), Dr. Masanobu Nanno (deputy director of Yakult Central Institute and director of Yakult Honsha European Research Center for Microbiology VOF), Dr. Masato Nagaoka (executive officer of Yakult Central Institute), Dr. Haruji Sawada (president of Yakult Honsha European Research Center for Microbiology VOF), Mr. Hiroyasu Matsubara and Dr. Takuya Takahashi (directors of Yakult Honsha European Research Center for Microbiology VOF), and Dr. Takashi Sato, Mr. Toru Iino, Mr. Yasuhisa Shimakawa and Dr. Akimitsu Takagi (my supervisors), as well, for their understanding and encouragement through my research activities. And I am also grateful to Dr. Bruno Pot, Dr. Bruno Biavati, Dr. Paola Mattarelli, Dr. Monica Modesto, Dr. Kenji Oishi, Dr. Eiji Ishikawa, Dr. Kazunori Matsuda, Dr. Agata Gawad, Dr. Hiroshi Makino, Dr. Takuya Akiyama, Ms. Yuko Kudo, Ms. Emiko Yanokura and my other colleagues of Culture Collection & Microbial Systematics Laboratory and Food Development Laboratory IV in Yakult Central Institute as well as Yakult Honsha European Research Center for Microbiology VOF for their assistance to my research.

Abstract in Japanese

微生物は地球上の様々な環境に分布し、その環境条件（例えば、pH、酸素濃度、栄養源）に適応して複雑な生態系を形成している。微生物生態学の目的は、微生物集団の構造と機能を説明する法則を見出すことにある。そのためには、対象とする生態系にどのような微生物が存在し、それらがどのように振舞うかを知る必要がある。近年、宿主と共生微生物との相互作用の重要性が認識されており、これらは 1 つの生態系の単位（ホロビオント）として取り扱われるべきであるという指摘がある。ヒトでは、特にその腸内に数百菌種におよぶ膨大な数の微生物を擁する複雑な生態系が構築されていることから、ヒトの健康を研究する上で腸内微生物の影響はもはや無視することはできない。ヒト腸内微生物生態に関し、特に宿主の健康・疾病との関連に注目した研究が近年活発に行われているが、その多くは属以上の分類階級に注目しており、特定の細菌種および細菌株の機能や、その挙動に関する知見は十分に整備されているとは言えない。すなわち、ヒト腸内における複雑な微生物生態系のさらなる理解には、より詳細な微生物の構成および個々の微生物の挙動について情報を蓄積する必要がある。加えて、微生物の遺伝学的・生理学的特徴の推定には、正しい分類同定が不可欠であり、そのためには、既存の分類体系の見直しおよび分類法の改善を継続的に行っていく必要がある。ヒト腸内微生物の中でも、*Bifidobacterium* 属および *Lactobacillus* 属細菌は、宿主にとって有用な構成微生物として知られており、プロバイオティクスとして商業的に利用される菌株も含まれる。そこで本研究では、ヒト腸内微生物のなかでも、機能的および商業的に重要な上記の細菌群に注目し、その分類と挙動を明らかにすることを目的とした。

本論文の第 1 章では背景および本研究の位置付けについて論じた。

また、第 2 章では、ヒト腸内の *Bifidobacterium* 属および *Lactobacillus* 属細菌の分類学的研究を行った。まず、ヒト腸内細菌の有用細菌群の 1 つとして知られる *Lactobacillus* 属細菌に着目し、ヒト腸内における新菌種の探索および新菌種提案を行った。ヤクルト本社中央研究所では、乳幼児から老人まで様々な年代の健常な日本人（61 名；年齢 0～101 歳）から分離した未同定の 328 細菌株を保有していた。上記細菌株を対象として

16S rRNA 遺伝子塩基配列に基づく再同定を行った結果、既存の細菌種とそれぞれ明確に区別される 2 細菌株（YIT 12363^{T} および YIT 12364^{T}）が見出された。さらに、各細菌株の生理学的・遺伝学的特徴に関する追加解析を行い、それぞれの近縁細菌種の基準株と比較した結果、上記の 2 細菌株はそれぞれ独立した細菌種に分類されることが確認された。上記を踏まえ、健常成人由来の YIT 12363^{T} および健常な百歳老人由来の YIT 12364^{T} をそれぞれ基準株として、*Lactobacillus saniviri* および *Lactobacillus senioris* を新細菌種として提案した。次に、ヒトおよび動物の主要腸内細菌種である *Bifidobacterium longum* について、亜種同定法の開発および既存の分類体系の再編提案を行った。*B. longum* に属する 3 つの亜種（subsp. *infantis*、subsp. *longum* および subsp. *suis*）は、元々はそれぞれ独立した細菌種として分類されていたが、長年に渡って蓄積された表現型および遺伝型の情報に基づき、2002 年に同一細菌種として現在の *B. longum* に再分類された。この 3 亜種は互いに非常に近縁であるため、それぞれを識別するためには、非常に時間と手間がかかることから、簡便迅速な識別方法が求められてきた。本報では、既存の手法により亜種レベルでの同定が完了している *B. longum* 25 細菌株を、Multi Locus Sequence Analysis 法、Multi Locus Sequence Typing 法および Amplified Fragment Length Polymorphism 法を用いて再分類した。その結果、いずれの手法においても 3 つの亜種に対応した集団が形成されることが認められ、これらの手法を用いることで簡便迅速かつ高感度に *B. longum* の細菌株を亜種レベルで識別できることを確認した。加えて、我々は *B. longum* subsp. *suis* に分類されている細菌群が 2 つの小集団に分かれること、それぞれの小集団はウレアーゼ活性の有無と一致することを明らかにした。このうちウレアーゼ活性を持たない *B. longum* subsp. *suis* 内の小集団は、豚の糞便から分離された細菌株で構成されていたことから、これらの菌株を *B. longum* subsp. *suillum* として新亜種提案した。

続いて、第 3 章では、*Bifidobacterium* 属細菌について、特にこの細菌群が最優勢となる乳児腸管における挙動に関する研究を行った。*Bifidobacterium* 属細菌の中でも、乳児腸管の優勢亜種として知られる *Bifidobacterium longum* subsp. *longum* に着目し、生後初期に定着した同亜種の細菌株が、その後同一被験児の腸内に長期間存在し続けるか検証した。ベルギーの乳児 12 名を対象として、生後半年まで（以下、出生期）の 7 点（胎

便、出生後 3 日、7 日、30 日、90 日および 180 日、または卒乳 1 週間後）および約 6 歳時点（以下、幼児期）に糞便を採取した。また、同亜種の細菌株については、出産前の母親の糞便または母乳からが生後初期の乳児腸管へ、同一の細菌株が伝播（母子伝播）することが報告されていた。そのため、本研究では長期定着と母子伝播との関連についても検証するため、上記 12 名の被験児の母親の出産前の糞便および出生後の母乳を採取した。各サンプルから分離した *B. longum* subsp. *longum* 462 細菌株について、7 種のハウスキーピング遺伝子を用いたタイピングに基づく細菌株同定を行った。その結果、上記細菌株は 73 種類の遺伝子型を示す異なる細菌株に分類された。さらに、各細菌株が分離されたサンプルを比較した結果、それぞれ異なる被験児由来の 3 細菌株が、出生期から幼児期までの間、同一被験児の腸内に定着し続けており、さらに、その内 1 細菌株は被験児の母親の母乳中にも存在したことが明らかになった。また、定量的 PCR を用いて、出生期および幼児期における各被験児の腸内 *Bifidobacterium* 属細菌の構成を解析した結果、上記 3 細菌株の長期定着細菌株は、そのほかの *Bifidobacterium* 属細菌群と共存し続けていたことが示唆された。本研究により、生後初期のヒト腸管への定着が、その後の長期的定着に重要であることが示された。

最後に、第 4 章では総括および全体を通した考察を行った。

以上、本研究では、ヒト腸内の *Bifidobacterium* 属および *Lactobacillus* 属細菌の分類および挙動についてさらなる知見を得ることができた。本研究で新たに提案した菌種および亜種により、これまでその存在が認識されていなかったこれらの微生物群が持つ生態学的機能に関する研究が、今後進むことが期待される。また、本研究で開発した *B. longum* の亜種分類法により、今後新たに分離培養される細菌株や、すでにライブラリ化されている細菌株の亜種同定を行うことで、同細菌種を対象とした研究が今後さらに発展していくであろう。さらに、*B. longum* subsp. *longum* の長期定着に関する研究では、生後初期期にヒト腸内に定着した *Bifidobacterium* 属の細菌株が、宿主の健康に長期的な効果を及ぼし得るという、生後初期のヒト腸内における微生物生態学の発展に貢献する重要な可能性が示された。

ヒト腸内微生物の機能については、近年環境汚染物質との相互作用およびそれが宿主

に与える影響について知見が蓄積しつつある。例えば、様々な環境汚染物質がヒト腸内微生物の構成や代謝活性を変化させる一方、ヒト腸内微生物に代謝されることでその毒性を変化させる事例が報告されている。環境汚染物質がヒトの健康に与える影響が世界的に懸念されるなか、ヒト腸内に存在する微生物のさらなる発見と理解によって、より正確にこれらのリスクを予測することが可能になるかもしれない。また、*Bifidobacterium* 属および *Lactobacillus* 属細菌の中には、環境汚染物質の毒性を緩和するものが報告されていることから、ヒト腸内におけるこれらの菌群の構成を変化させることにより、同様の毒性低減効果が実現できるかもしれない。さらに、一部の *Bifidobacterium* 属および *Lactobacillus* 属の菌株には、有害重金属のような環境汚染物質に対する吸着活性が報告されている。そのため、実際に効果をもたらしている菌体成分が特定されれば、それらの環境浄化へ応用した技術の開発も期待できる。

Bifidobacterium 属および *Lactobacillus* 属細菌は、ヒト腸内における有用菌群として今後さらにその重要性が増すことが予測され、本論文は、当該研究同分野の一助となるものである。

Supplementary Data

Table S1. Detailed information of the subjects participating in the follow-up study.

Subject ID	Gender[a]	Delivery mode[b]	1st mothers prenatal fecal sample[c] (days)	2nd mothers prenatal fecal sample[c] (days)	After weaning[c,d] (days)	Age at follow-up (years)	Isolation of *B. longum* subsp. *longum* from subject's feces both in early infancy and childhood
4	M	C	−42	−34	133	6.9	●
12	F	N	−59	−47	> 180	6.7	●
13	F	N	−31	−24	120	6.7	●
16	F	N	−38	−31	> 180	6.8	●
18	M	N	−21	−9	136	6.5	●
20	M	N	−47	−40	130	6.4	−
26	F	N	−36	−29	106	6.7	−
28	M	N	−43	−35	151	4.3	−
30	M	C	−25	−18	130	6.7	●
37	F	N	−32	−25	> 180	6.6	−
38	M	N	−49	−42	99	6.3	●
44	M	N	−19	−12	110	6.6	●
48	M	C	−49	−37	124	6.3	●
49	F	C	−35	−28	157	6.4	−
52	F	N	−15	−4	> 180	6.3	−
53	M	N	−39	−31	> 180	6.4	−
57	F	N	−38	−31	141	6.4	−
61	F	N	−26	−17	167	6.6	−
62	M	N	−21	−11	126	6.3	−
63	M	N	−30	−24	> 180	6.6	−
68	M	N	−41	−33	150	6.6	−
71	F	C	−38	−31	160	6.3	−
84	M	N	−50	−44	159	6.2	−
86	F	N	−33	−26	158	6.6	−
94	F	N	−36	−28	165	6.3	−
95	F	N	−49	−42	175	6.4	−
96	F	N	−31	−25	> 180	6.4	−
100	M	N	−49	−42	138	6.3	−
106	M	C	−52	−45	130	6.2	−
107	M	N	−48	−41	> 180	6.2	−
108A (Twin)	F	C	−57	−50	> 180	6.1	−
108B (Twin)	F	C	−57	−50	> 180	6.1	−
109	F	N	−52	−45	148	6.3	−
117	M	N	−41	−34	129	6.2	−
120	M	N	−46	−38	> 180	7.4	−
122	F	N	−35	−28	96	6.1	−
123	M	C	−56	−50	131	5.8	−
126	F	N	−37	−31	134	6.0	−
129	F	N	−57	−50	128	5.9	●
130	M	N	−55	−47	124	6.1	●
131	M	C	−45	−38	155	6.1	−
133	F	C	−51	−43	147	5.9	−
134	M	N	−54	−44	162	5.8	●
136	F	N	−34	−25	148	5.8	−
137	M	N	−46	−38	147	5.9	−
138	M	N	−36	−30	> 180	8.7	−
139	F	N	−27	−19	172	6.1	−
140	M	N	−48	−40	133	5.8	−
141	F	N	−50	−43	161	5.8	−

[a] M, male; F, female. [b] N, natural delivery; C, cesarean section.

[c] Sample collection date calculated from delivery date. [d] One week after the introduction of solids.

Table S2. Count of *B. longum* subsp. *longum* isolates obtained from the samples collected from the 12 subjects and their mothers.

Subject ID	Count of isolates[a]											
	Mother's perinatal samples				Subject's samples							
	Prenatal feces		Postnatal breast milk		Feces in early infancy							Feces in childhood
	1st Sample	2nd Sample	7 days	30 days	Meconium	3 days	7 days	30 days	90 days	After Weaning[b]	180 days	≈ 6 years
4	3	4	−	−	−	−	−	−	−	3	−	4
12	11	7	−	−	−	3	8	4	−	ns	−	3
13	9	11	−	−	−	−	−	−	−	3	4	9
16	12	9	−	−	4	26	22	20	24	ns	5	7
18	9	8	1	−	−	6	6	6	5	−	−	3
30	4	5	−	−	−	ns	−	−	3	2	−	1
38	11	12	−	−	−	11	12	12	−	−	−	2
44	3	−	−	1	−	−	2	−	−	4	−	1
48	−	7	−	1	−	ns	−	6	−	−	−	2
129	−	−	12	8	3	3	3	12	−	−	−	3
130	5	7	−	−	−	3	2	−	−	−	−	7
134	4	−	−	9	−	4	6	6	−	−	−	4

[a] ns, no sample was collected; −, no isolate was obtained. [b] One week after the introduction of solids.

Table S3. Detailed information of the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum*.

Strain ID	Subject ID	Isolated sample type	Detection time point[a]	ST	Allelic numbers							Reference fo sequences
					clpC	*dnaG*	*dnaJ*	*fusA*	*gyrB*	*purF*	*rpoB*	
31	4	Mother feces (prenatal)	1st sample	4-A	1	20	1	3	1	4	3	This study
35	4	Mother feces (prenatal)	1st sample	4-B	1	21	1	1	12	1	14	This study
1800	4	Subject feces (early infancy)	After weaning	4-C	1	7	1	6	1	6	11	This study
4-2	4	Subject feces (childhood)	≈ 6 years	4-D	1	3	1	1	1	4	3	This study
4-4	4	Subject feces (childhood)	≈ 6 years	4-E	1	1	1	3	7	2	3	This study
4-6	4	Subject feces (childhood)	≈ 6 years	4-F	1	1	1	3	1	2	3	This study
237	12	Mother feces (prenatal)	1st sample	12-A	1	5	1	1	1	2	14	This study
818	12	Subject feces (early infancy)	3 days	12-A	1	5	1	1	1	2	14	This study
238	12	Mother feces (prenatal)	1st sample	12-B	6	12	1	10	10	1	3	This study
856	12	Subject feces (early infancy)	7 days	12-B	6	12	1	10	10	1	3	This study
239	12	Mother feces (prenatal)	1st sample	12-C	6	12	1	1	10	1	15	This study
241	12	Mother feces (prenatal)	1st sample	12-D	1	12	1	10	3	1	3	This study
244	12	Mother feces (prenatal)	1st sample	12-E	1	5	1	10	1	1	15	This study
246	12	Mother feces (prenatal)	1st sample	12-F	6	5	1	1	1	11	15	This study
248	12	Mother feces (prenatal)	1st sample	12-G	6	12	1	1	1	2	15	This study
334	12	Mother feces (prenatal)	2nd sample	12-H	6	12	1	1	10	2	3	This study
336	12	Mother feces (prenatal)	2nd sample	12-I	1	4	4	3	1	2	5	This study
339	12	Mother feces (prenatal)	2nd sample	12-J	6	12	1	3	10	1	15	This study
12-4	12	Subject feces (childhood)	≈ 6 years	12-K	1	16	6	11	3	2	2	This study
12-5	12	Subject feces (childhood)	≈ 6 years	12-L	1	17	6	12	3	4	8	This study
1784	13	Subject feces (early infancy)	After weaning	13-A	1	1	1	3	1	5	10	This study
13-T1	13	Subject feces (childhood)	≈ 6 years	13-A	1	1	1	3	1	5	10	This study
139	13	Mother feces (prenatal)	1st sample	13-B	3	3	1	1	5	2	7	This study
141	13	Mother feces (prenatal)	1st sample	13-C	1	3	3	1	6	2	7	This study
144	13	Mother feces (prenatal)	1st sample	13-D	5	6	1	3	1	2	8	This study
220	13	Mother feces (prenatal)	2nd sample	13-E	1	9	4	8	6	2	10	This study
1847	13	Subject feces (early infancy)	180 days	13-F	5	1	1	3	1	5	10	This study
1848	13	Subject feces (early infancy)	180 days	13-G	5	1	1	3	1	5	13	This study
13-1	13	Subject feces (childhood)	≈ 6 years	13-H	1	3	2	1	1	2	16	This study
13-T2	13	Subject feces (childhood)	≈ 6 years	13-I	1	1	4	3	1	5	10	This study
281	16	Mother feces (prenatal)	1st sample	16-A	1	4	1	3	3	1	5	Makino, et al. 2013
401	16	Subject feces (early infancy)	Meconium	16-A	1	4	1	3	3	1	5	Makino, et al. 2013
286	16	Mother feces (prenatal)	1st sample	16-B	1	14	1	3	3	4	5	Makino, et al. 2013
435	16	Subject feces (early infancy)	3 days	16-B	1	14	1	3	3	4	5	Makino, et al. 2013
297	16	Mother feces (prenatal)	1st sample	16-C	1	15	1	3	4	1	5	Makino, et al. 2013
441	16	Subject feces (early infancy)	3 days	16-C	1	15	1	3	4	1	5	Makino, et al. 2013
326	16	Mother feces (prenatal)	2nd sample	16-D	1	15	1	4	4	3	5	Makino, et al. 2013
520	16	Subject feces (early infancy)	7 days	16-D	1	15	1	4	4	3	5	This study
298	16	Mother feces (prenatal)	1st sample	16-E	8	4	1	3	4	1	5	This study
296	16	Mother feces (prenatal)	1st sample	16-F	1	4	1	3	4	1	4	This study
485	16	Subject feces (early infancy)	7 days	16-G	1	15	1	4	4	4	4	This study
521	16	Subject feces (early infancy)	7 days	16-H	1	15	1	4	4	4	5	This study
527	16	Subject feces (early infancy)	7 days	16-I	1	15	1	4	3	4	4	This study
724	16	Subject feces (early infancy)	30 days	16-J	8	15	1	3	3	4	5	This study
1305	16	Subject feces (early infancy)	90 days	16-K	1	4	1	3	3	1	4	This study
1314	16	Subject feces (early infancy)	90 days	16-L	1	4	1	4	4	3	5	This study
16-1	16	Subject feces (childhood)	≈ 6 years	16-M	1	14	1	3	3	1	5	This study
16-T1	16	Subject feces (childhood)	≈ 6 years	16-N	1	15	1	4	4	1	5	This study
865	18	Mother feces (prenatal)	1st sample	18-A	1	2	2	2	2	2	2	This study
1120	18	Subject feces (early infancy)	3 days	18-A	1	2	2	2	2	2	2	This study
1227	18	Breast milk	7 days	18-A	1	2	2	2	2	2	2	This study
18-1	18	Subject feces (childhood)	≈ 6 years	18-B	1	5	1	3	3	2	16	This study
1521	30	Subject feces (early infancy)	90 days	30-A	1	1	1	5	7	1	3	Makino, et al. 2013
30-1	30	Subject feces (childhood)	≈ 6 years	30-A	1	1	1	5	7	1	3	This study
630	30	Mother feces (prenatal)	1st sample	30-B	1	1	1	1	7	1	13	Makino, et al. 2013
631	30	Mother feces (prenatal)	1st sample	30-C	1	3	1	3	7	12	13	Makino, et al. 2013
1082	38	Mother feces (prenatal)	1st sample	38-A	1	1	1	1	1	1	1	This study
1699	38	Subject feces (early infancy)	3 days	38-A	1	1	1	1	1	1	1	This study
1835	38	Mother feces (prenatal)	1st sample	38-B	1	3	1	3	8	4	3	This study
1688	38	Subject feces (early infancy)	7 days	38-B	1	3	1	3	8	4	3	This study
1078	38	Mother feces (prenatal)	1st sample	38-C	1	4	8	3	1	2	9	This study
1687	38	Subject feces (early infancy)	7 days	38-D	1	4	1	3	1	2	9	This study
38-1	38	Subject feces (childhood)	≈ 6 years	38-E	1	3	1	1	3	2	19	This study
977	44	Mother feces (prenatal)	1st sample	44-A	2	3	1	1	1	1	3	This study
1268	44	Subject feces (early infancy)	7 days	44-A	2	3	1	1	1	1	3	This study
1426	44	Breast milk	30 days	44-B	4	1	1	3	6	2	3	This study
1854	44	Subject feces (early infancy)	After weaning	44-B	4	1	1	3	6	2	3	This study
44-1	44	Subject feces (childhood)	≈ 6 years	44-B	4	1	1	3	6	2	3	This study
1620	48	Breast milk	30 days	48-A	4	8	1	7	1	4	12	This study
1841	48	Subject feces (early infancy)	30 days	48-A	4	8	1	7	1	4	12	Makino, et al. 2013
1362	48	Mother feces (prenatal)	2nd sample	48-B	1	5	1	3	1	4	6	Makino, et al. 2013
1844	48	Subject feces (early infancy)	30 days	48-C	4	8	1	7	1	2	12	Makino, et al. 2013
48-1	48	Subject feces (childhood)	≈ 6 years	48-D	1	3	1	14	3	4	12	This study
48-2	48	Subject feces (childhood)	≈ 6 years	48-E	1	3	8	3	3	14	12	This study
2285	129	Subject feces (early infancy)	Meconium	129-A	1	10	1	1	1	2	2	This study
2220	129	Breast milk	7 days	129-A	1	10	1	1	1	2	2	This study
129-4	129	Subject feces (childhood)	≈ 6 years	129-B	1	3	1	9	1	2	3	This study
2296	130	Mother feces (prenatal)	1st sample	130-A	1	4	1	9	9	8	13	This study
2317	130	Subject feces (early infancy)	3 days	130-A	1	4	1	9	9	8	13	This study
2297	130	Mother feces (prenatal)	1st sample	130-B	1	11	1	1	9	7	2	This study
2316	130	Subject feces (early infancy)	3 days	130-B	1	11	1	1	9	7	2	This study
2302	130	Mother feces (prenatal)	1st sample	130-C	1	4	1	9	3	9	13	This study
2315	130	Subject feces (early infancy)	3 days	130-C	1	4	1	9	3	9	13	This study
2295	130	Mother feces (prenatal)	1st sample	130-D	1	11	5	1	9	7	2	This study
2308	130	Mother feces (prenatal)	2nd sample	130-E	1	11	5	1	3	7	2	This study
130-1	130	Subject feces (childhood)	≈ 6 years	130-F	1	18	1	13	3	4	17	This study
130-3	130	Subject feces (childhood)	≈ 6 years	130-G	1	18	1	13	3	2	17	This study
130-5	130	Subject feces (childhood)	≈ 6 years	130-H	1	18	1	13	6	4	17	This study
130-6	130	Subject feces (childhood)	≈ 6 years	130-I	1	18	7	13	6	4	17	This study
2476	134	Mother feces (prenatal)	1st sample	134-A	7	3	1	1	11	10	14	Makino, et al. 2013
2488	134	Subject feces (early infancy)	3 days	134-A	7	3	1	1	11	10	14	Makino, et al. 2013
2452	134	Breast milk	30 days	134-A	7	3	1	1	11	10	14	This study
2493	134	Subject feces (early infancy)	3 days	134-B	1	13	2	3	3	2	1	This study
2459	134	Breast milk	30 days	134-B	1	13	2	3	3	2	1	This study
134-1	134	Subject feces (childhood)	≈ 6 years	134-C	1	19	1	3	1	13	18	This study
134-4	134	Subject feces (childhood)	≈ 6 years	134-D	7	5	1	3	7	2	3	This study
B. longum subsp. *infantis* (JCM 1222^T)	—	—	—	INF	9	22	9	15	13	3	20	Fukuda, et al. 2011
B. longum subsp. *longum* (JCM 1217^T)	—	—	—	LON	5	14	1	3	14	2	3	Fukuda, et al. 2011

[a] Calculated from the delivery date; After weaning, one week after the introduction of solids.

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum*.

Strain ID	ST	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*clpC*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*
	Site number[a]	27	78	93	102	153	156	201	210	225	276	312	315	357	360	431	6	9	15	21	24	30	39	53	54	63	64	69	84	85	86
31	4-A	T	C	T	C	C	T	C	G	C	T	T	C	A	C	C	C	C	C	G	G	C	G	T	T	T	C	T	C	A	G
35	4-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	A	·	·	C	·	·	·	·	T	·	·
1800	4-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
4-2	4-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
4-4	4-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
237	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
818	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
238	12-B	C	·	C	T	·	·	·	C	·	C	·	·	·	·	T	·	T	·	·	·	G	·	C	C	G	T	A	·	·	C
856	12-B	C	·	C	T	·	·	·	C	·	C	·	·	·	·	T	·	T	·	·	·	G	·	C	C	G	T	A	·	·	C
239	12-C	C	·	C	T	·	·	·	C	·	C	·	·	·	·	T	·	T	·	·	·	G	·	C	C	G	T	A	·	·	C
241	12-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	G	·	C	C	G	T	A	·	·	C
244	12-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
246	12-F	C	·	C	T	·	·	·	C	·	C	·	·	·	·	T	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
248	12-G	C	·	C	T	·	·	·	C	·	C	·	·	·	·	T	·	T	·	·	·	G	·	C	C	G	T	A	·	·	C
334	12-H	C	·	C	T	·	·	·	C	·	C	·	·	·	·	T	·	T	·	·	·	G	·	C	C	G	T	A	·	·	C
336	12-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
339	12-J	C	·	C	T	·	·	·	C	·	C	·	·	·	·	T	·	T	·	·	·	G	·	C	C	G	T	A	·	·	C
12-4	12-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	G	·	C	·	·	·	·	·	·	·
12-5	12-L	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
139	13-B	·	·	·	·	·	·	T	·	T	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
141	13-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
144	13-D	·	·	·	·	·	·	·	·	·	C	·	T	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
220	13-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	G	·	C	C	G	T	A	·	G	·
1847	13-F	·	·	·	·	·	·	·	·	·	C	·	T	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	C	·	T	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
13-1	13-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
281	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
401	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
286	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	C	·	·	·	·	·	·	·
435	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	C	·	·	·	·	·	·	·
297	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
441	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
326	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
520	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
298	16-E	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
296	16-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
485	16-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
521	16-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
527	16-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
724	16-J	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
1305	16-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
1314	16-L	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
16-1	16-M	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	C	·	·	·	·	·	·	·
16-T1	16-N	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	·	C
865	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	·	·	·	·	·	·	·
1120	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	·	·	·	·	·	·	·
1227	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	·	·	·	·	·	·	·
18-1	18-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
1521	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
30-1	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
630	30-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
631	30-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
1082	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
1699	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
1835	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
1688	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
1078	38-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
1687	38-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
38-1	38-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
977	44-A	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
1268	44-A	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
1426	44-B	·	·	·	·	·	·	·	·	·	·	·	·	G	T	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
1854	44-B	·	·	·	·	·	·	·	·	·	·	·	·	G	T	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
44-1	44-B	·	·	·	·	·	·	·	·	·	·	·	·	G	T	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
1620	48-A	·	·	·	·	·	·	·	·	·	·	·	·	G	T	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
1841	48-A	·	·	·	·	·	·	·	·	·	·	·	·	G	T	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
1362	48-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
1844	48-C	·	·	·	·	·	·	·	·	·	·	·	·	G	T	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
48-1	48-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
48-2	48-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
2285	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	G	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	G	·	·	·	·	·
129-4	129-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
2296	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
2317	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
2297	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
2316	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
2302	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
2315	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	·	C
2295	130-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
2308	130-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
130-1	130-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	G	·
130-3	130-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	G	·
130-5	130-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	G	·
130-6	130-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	C	C	G	T	A	·	G	·
2476	134-A	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
2488	134-A	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
2452	134-A	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	A	·	G	·	C	C	G	T	A	·	G	·
2493	134-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	G	T	A	·	G	·
2459	134-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	G	T	A	·	G	·
134-1	134-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	C	C	G	T	A	·	G	·
134-4	134-D	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222[T])	INF	C	·	C	T	T	A	·	·	·	C	·	·	·	·	·	·	·	A	·	·	·	A	C	C	G	T	A	·	G	·
B. longum subsp. *longum* (JCM 1217[T])	LON	·	·	·	·	·	·	·	·	·	C	·	T	·	·	·	T	·	·	·	·	·	·	C	·	·	·	·	·	·	·

For each gene, all discovered alleles were compared, but only polymorphic sites are shown. Numbering starts at the beginning of the aligned sequence portion of each gene.

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued 1).

Strain ID	ST	Gene names *dnaG* Site number[a] 87	*dnaG* 90	*dnaG* 92	*dnaG* 93	*dnaG* 96	*dnaG* 102	*dnaG* 105	*dnaG* 108	*dnaG* 109	*dnaG* 110	*dnaG* 115	*dnaG* 117	*dnaG* 122	*dnaG* 124	*dnaG* 150	*dnaG* 156	*dnaG* 162	*dnaG* 165	*dnaG* 171	*dnaG* 174	*dnaG* 180	*dnaG* 189	*dnaG* 201	*dnaG* 207	*dnaG* 210	*dnaG* 213	*dnaG* 222	*dnaG* 223	*dnaG* 225	*dnaG* 226
31	4-A	T	C	A	C	T	G	G	G	G	T	G	C	T	T	C	C	A	A	G	T	C	G	C	T	G	T	A	T	G	T
35	4-B	·	·	·	·	·	·	·	T	C	C	T	G	C	C	·	T	C	G	·	C	·	A	T	C	·	C	·	·	A	C
1800	4-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-2	4-D	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
4-4	4-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
237	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
818	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
238	12-B	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	C	A	G	G	C	A	C
856	12-B	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	C	A	G	G	C	A	C
239	12-C	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	C	A	G	G	C	A	C
241	12-D	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	C	A	G	G	C	A	C
244	12-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
246	12-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
248	12-G	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	C	A	G	G	C	A	C
334	12-H	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	C	A	G	G	C	A	C
336	12-I	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
339	12-J	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	C	A	G	G	C	A	C
12-4	12-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12-5	12-L	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
139	13-B	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
141	13-C	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
144	13-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
220	13-E	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	·	T	C	A	G	G	C	A	C
1847	13-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-1	13-H	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
281	16-A	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
401	16-A	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
286	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
435	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
297	16-C	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
441	16-C	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
326	16-D	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
520	16-D	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
298	16-E	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
296	16-F	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
485	16-G	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
521	16-H	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
527	16-I	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
724	16-J	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1305	16-K	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1314	16-L	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-1	16-M	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-T1	16-N	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
865	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1120	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1227	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
18-1	18-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1521	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
30-1	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
630	30-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
631	30-C	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
1082	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1699	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1835	38-B	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
1688	38-B	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
1078	38-C	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1687	38-D	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
38-1	38-E	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
977	44-A	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
1268	44-A	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
1426	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1854	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
44-1	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1620	48-A	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
1841	48-A	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
1362	48-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1844	48-C	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
48-1	48-D	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
48-2	48-E	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
2285	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
129-4	129-B	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
2296	130-A	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2317	130-A	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2297	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·
2316	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·
2302	130-C	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2315	130-C	C	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2295	130-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·
2308	130-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·
130-1	130-F	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	G	C	A	C
130-3	130-G	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	G	C	A	C
130-5	130-H	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	G	C	A	C
130-6	130-I	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	G	C	A	C
2476	134-A	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
2488	134-A	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
2452	134-A	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
2493	134-B	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
2459	134-B	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	A	T	C	A	G	·	·	·	·
134-1	134-C	·	T	·	T	C	A	A	T	C	C	T	G	·	C	T	G	C	G	A	C	·	·	T	C	A	G	G	C	A	C
134-4	134-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222[T])	INF	·	T	·	T	C	A	A	T	C	C	T	G	C	C	T	G	C	G	A	C	·	A	T	C	A	G	·	C	A	C
B. longum subsp. *longum* (JCM 1217[T])	LON	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued 2).

Strain ID	ST	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaG*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*	*dnaJ*
	Site number[a]	246	255	279	285	288	294	297	300	303	13	14	29	32	33	38	41	53	65	71	74	80	85	124	143	155	158	183	188	189	190
31	4-A	C	T	A	C	C	A	G	G	A	A	T	C	C	A	T	T	C	C	T	C	C	G	A	C	T	G	G	A	T	C
35	4-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1800	4-C	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-2	4-D	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-4	4-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
237	12-A	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
818	12-A	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
238	12-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
856	12-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
239	12-C	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
241	12-D	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
244	12-E	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
246	12-F	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
248	12-G	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
334	12-H	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
336	12-I	·	·	G	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
339	12-J	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12-4	12-K	·	·	·	·	·	·	·	·	·	·	C	T	T	·	C	C	·	T	C	T	T	·	·	·	·	·	·	·	·	·
12-5	12-L	·	·	·	·	·	·	·	·	·	·	C	T	T	·	C	C	·	T	C	T	T	·	·	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
139	13-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
141	13-C	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·
144	13-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
220	13-E	·	C	G	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1847	13-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-1	13-H	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	A	·	·	·	·
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
281	16-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
401	16-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
286	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
435	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
297	16-C	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
441	16-C	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
326	16-D	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
520	16-D	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
298	16-E	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
296	16-F	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
485	16-G	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
521	16-H	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
527	16-I	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
724	16-J	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1305	16-K	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1314	16-L	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-1	16-M	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-T1	16-N	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
865	18-A	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	A	·	·	·	·
1120	18-A	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	A	·	·	·	·
1227	18-A	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	A	·	·	·	·
18-1	18-B	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1521	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
30-1	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
630	30-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
631	30-C	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1082	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1699	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1835	38-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1688	38-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1078	38-C	·	·	G	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1687	38-D	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
38-1	38-E	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
977	44-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1268	44-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1426	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1854	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
44-1	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1620	48-A	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1841	48-A	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1362	48-B	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1844	48-C	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
48-1	48-D	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
48-2	48-E	·	·	G	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2285	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
129-4	129-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2296	130-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2317	130-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2297	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2316	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2302	130-C	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2315	130-C	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2295	130-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·
2308	130-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·
130-1	130-F	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-3	130-G	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-5	130-H	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-6	130-I	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·
2476	134-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2488	134-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2452	134-A	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2493	134-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	A	·	·	·	·
2459	134-B	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	A	·	·	·	·
134-1	134-C	·	C	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-4	134-D	·	·	·	G	T	G	·	A	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222^T)	INF	·	C	G	·	·	·	A	A	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	C	A	A	G	C	A
B. longum subsp. *longum* (JCM 1217^T)	LON	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued-3).

Strain ID	ST	Gene names *dnaJ*	*dnaJ*	*dnaJ*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*
		Site number[a] 192	206	284	2	5	6	24	32	35	36	53	59	65	71	83	86	101	113	131	134	137	142	146	161	167	194	212	233	266	275
31	4-A	C	C	T	C	T	A	T	C	T	T	T	T	C	T	T	C	C	C	T	G	C	A	C	T	G	T	C	G	T	T
35	4-B	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
1800	4-C	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-2	4-D	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
4-4	4-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
237	12-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
818	12-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
238	12-B	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	T	·	·	A	·	·	·	·	·	·	·	·	·	·
856	12-B	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	T	·	·	A	·	·	·	·	·	·	·	·	·	·
239	12-C	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
241	12-D	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	T	·	·	A	·	·	·	·	·	·	·	·	·	·
244	12-E	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	T	·	·	A	·	·	·	·	·	·	·	·	·	·
246	12-F	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
248	12-G	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
334	12-H	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
336	12-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
339	12-J	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12-4	12-K	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
12-5	12-L	·	·	·	·	·	·	C	·	·	·	C	C	·	C	C	T	·	·	C	·	·	·	T	C	·	C	A	·	C	C
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
139	13-B	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
141	13-C	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
144	13-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
220	13-E	·	·	·	·	·	·	C	·	·	·	C	·	·	·	C	·	·	T	C	·	·	·	·	C	A	·	G	·	·	C
1847	13-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-1	13-H	·	·	C	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
281	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
401	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
286	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
435	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
297	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
441	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
326	16-D	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	A	·	·
520	16-D	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	A	·	·
298	16-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
296	16-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
485	16-G	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	A	·	·
521	16-H	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	A	·	·
527	16-I	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	A	·	·
724	16-J	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1305	16-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1314	16-L	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	A	·	·
16-1	16-M	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-T1	16-N	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	A	·	·
865	18-A	·	·	C	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
1120	18-A	·	·	C	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
1227	18-A	·	·	C	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
18-1	18-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1521	30-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	G	·	·	·	·	·	·	·	·
30-1	30-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	G	·	·	·	·	·	·	·	·
630	30-B	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
631	30-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1082	38-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
1699	38-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
1835	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1688	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1078	38-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1687	38-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
38-1	38-E	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
977	44-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
1268	44-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
1426	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1854	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
44-1	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1620	48-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	C	·	·	·	·	·	·
1841	48-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	C	·	·	·	·	·	·
1362	48-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1844	48-C	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	C	·	·	·	·	·	·
48-1	48-D	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
48-2	48-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2285	129-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2220	129-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
129-4	129-B	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2296	130-A	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2317	130-A	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2297	130-B	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2316	130-B	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2302	130-C	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2315	130-C	·	·	·	T	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2295	130-D	·	T	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2308	130-E	·	T	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
130-1	130-F	·	·	·	T	·	·	C	·	·	·	·	·	T	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
130-3	130-G	·	·	·	T	·	·	C	·	·	·	·	·	T	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
130-5	130-H	·	·	·	T	·	·	C	·	·	·	·	·	T	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
130-6	130-I	·	·	·	T	·	·	C	·	·	·	·	·	T	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2476	134-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2488	134-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2452	134-A	·	·	·	T	·	·	C	T	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·
2493	134-B	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2459	134-B	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-1	134-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-4	134-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222[T])	INF	G	·	·	·	C	G	C	·	C	G	C	C	·	C	C	·	·	·	C	·	T	·	T	C	·	C	A	·	C	C
B. longum subsp. *longum* (JCM 1217[T])	LON	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued-4).

Strain ID	ST	Gene names: *fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*fusA*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*
	Site number[a]	306	314	317	359	362	368	380	383	386	416	422	437	443	444	446	456	468	473	494	15	27	33	39	60	68	73	78	93	96	102
31	4-A	A	C	T	C	T	T	T	T	A	T	T	C	T	C	C	T	G	T	C	C	T	T	T	T	C	A	C	G	T	C
35	4-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·
1800	4-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-2	4-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-4	4-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
4-6	4-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
237	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
818	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
238	12-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·
856	12-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·
239	12-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·
241	12-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
244	12-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
246	12-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
248	12-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
334	12-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·
336	12-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
339	12-J	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·
12-4	12-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12-5	12-L	G	T	·	T	C	C	·	C	·	C	C	·	·	·	·	A	·	C	·	·	·	·	·	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
139	13-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·
141	13-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
144	13-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
220	13-E	·	·	·	·	C	·	C	G	G	·	·	·	C	A	G	A	T	·	·	·	·	C	·	C	·	·	·	·	·	T
1847	13-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-1	13-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
281	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
401	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
286	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
435	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
297	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
441	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
326	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
520	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
298	16-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
296	16-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
485	16-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
521	16-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
527	16-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
724	16-J	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1305	16-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1314	16-L	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
16-1	16-M	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-T1	16-N	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	C	·	G	T	A	C	T
865	18-A	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1120	18-A	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1227	18-A	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
18-1	18-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1521	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
30-1	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
630	30-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
631	30-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
1082	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1699	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1835	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·
1688	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·
1078	38-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1687	38-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
38-1	38-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
977	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1268	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1426	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
1854	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
44-1	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
1620	48-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1841	48-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1362	48-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1844	48-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
48-1	48-D	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
48-2	48-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2285	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
129-4	129-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2296	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2317	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2297	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2316	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2302	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2315	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2295	130-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2308	130-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-1	130-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-3	130-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-5	130-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
130-6	130-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
2476	134-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
2488	134-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
2452	134-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
2493	134-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2459	134-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-1	134-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-4	134-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	·	T
B. longum subsp. *infantis* (JCM 1222[T])	INF	G	·	C	·	C	C	C	C	G	C	C	·	·	·	·	A	T	·	T	·	C	C	C	·	·	·	·	·	C	·
B. longum subsp. *longum* (JCM 1217[T])	LON	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued-5).

Strain ID	ST	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*gyrB*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*
Site number[a]		117	150	156	159	165	198	210	216	240	264	270	273	296	381	390	6	30	42	48	49	51	57	66	102	108	117	120	123	129	135
31	4-A	C	T	G	T	C	C	C	A	G	C	C	T	C	T	C	C	G	T	C	G	C	T	C	C	C	C	C	T	T	C
35	4-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1800	4-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
4-2	4-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-4	4-E	·	C	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
237	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
818	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
238	12-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
856	12-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
239	12-C	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
241	12-D	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
244	12-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
246	12-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
248	12-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
334	12-H	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
336	12-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
339	12-J	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
12-4	12-K	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
12-5	12-L	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
139	13-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
141	13-C	·	C	·	·	T	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
144	13-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
220	13-E	·	C	·	·	T	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1847	13-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
13-1	13-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
281	16-A	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
401	16-A	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
286	16-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
435	16-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
297	16-C	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
441	16-C	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
326	16-D	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
520	16-D	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
298	16-E	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
296	16-F	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
485	16-G	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
521	16-H	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
527	16-I	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
724	16-J	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1305	16-K	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1314	16-L	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
16-1	16-M	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
16-T1	16-N	T	C	A	C	·	G	G	G	A	·	·	·	·	·	T	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
865	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1120	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1227	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
18-1	18-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1521	30-A	·	C	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
30-1	30-A	·	C	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
630	30-B	·	C	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
631	30-C	·	C	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	T	T	A	C	T	T	·	·	·	·	·	·
1082	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1699	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1835	38-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1688	38-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1078	38-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1687	38-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
38-1	38-E	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
977	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1268	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1426	44-B	·	C	·	·	T	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1854	44-B	·	C	·	·	T	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
44-1	44-B	·	C	·	·	T	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
1620	48-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1841	48-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1362	48-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1844	48-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
48-1	48-D	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
48-2	48-E	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	T	·	C	C	G
2285	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
129-4	129-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
2296	130-A	·	·	·	·	·	·	·	·	·	T	·	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2317	130-A	·	·	·	·	·	·	·	·	·	T	·	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2297	130-B	·	·	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	G	·	·	·
2316	130-B	·	·	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	G	·	·	·
2302	130-C	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	·	·	·	·	·	·	·
2315	130-C	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	·	·	·	·	·	·	·
2295	130-D	·	·	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	G	·	·	·
2308	130-E	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	G	·	·	·
130-1	130-F	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-3	130-G	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
130-5	130-H	·	C	·	·	T	·	·	·	·	T	·	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-6	130-I	·	C	·	·	T	·	·	·	·	T	·	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2476	134-A	·	C	·	·	T	·	·	·	·	T	·	·	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
2488	134-A	·	C	·	·	T	·	·	·	·	T	·	·	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
2452	134-A	·	C	·	·	T	·	·	·	·	T	·	·	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
2493	134-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
2459	134-B	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
134-1	134-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-4	134-D	·	C	·	·	·	·	·	·	·	T	·	C	·	C	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222^T)	INF	·	C	A	C	·	G	G	G	A	·	G	·	·	C	·	T	A	C	·	T	A	C	T	T	·	·	·	·	·	·
B. longum subsp. *longum* (JCM 1217^T)	LON	·	C	·	·	·	·	·	·	·	·	·	·	T	·	·	·	A	C	·	T	A	C	T	T	·	·	·	·	·	·

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued-6).

Strain ID	ST	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*purF*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*
	Site number[a]	141	175	180	201	208	210	211	225	258	312	316	363	402	403	9	10	11	12	15	25	26	30	36	42	45	46	47	48	54	57
31	4-A	T	G	T	G	A	T	T	C	G	C	A	C	G	C	C	A	A	C	A	G	C	C	G	C	A	A	A	G	C	T
35	4-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1800	4-C	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-2	4-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-4	4-E	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
237	12-A	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
818	12-A	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
238	12-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
856	12-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
239	12-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
241	12-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
244	12-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
246	12-F	·	·	·	·	·	·	·	·	A	·	-	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
248	12-G	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
334	12-H	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
336	12-I	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
339	12-J	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12-4	12-K	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12-5	12-L	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
139	13-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
141	13-C	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
144	13-D	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
220	13-E	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1847	13-F	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
13-1	13-H	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
281	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
401	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
286	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
435	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
297	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
441	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
326	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
520	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
298	16-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
296	16-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
485	16-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
521	16-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
527	16-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
724	16-J	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1305	16-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
1314	16-L	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-1	16-M	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
16-T1	16-N	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
865	18-A	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1120	18-A	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1227	18-A	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
18-1	18-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
1521	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
30-1	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
630	30-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
631	30-C	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1082	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1699	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1835	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1688	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1078	38-C	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1687	38-D	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
38-1	38-E	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
977	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1268	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1426	44-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1854	44-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
44-1	44-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1620	48-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
1841	48-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
1362	48-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	T	C	·	C	A	A	T	A	T	G	G	C	T	·	C
1844	48-C	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
48-1	48-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
48-2	48-E	C	A	C	C	G	C	C	·	·	·	·	·	·	·	·	T	C	·	C	A	A	·	A	T	G	G	C	T	·	C
2285	129-A	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
129-4	129-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2296	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2317	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2297	130-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2316	130-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2302	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2315	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2295	130-D	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2308	130-E	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-1	130-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-3	130-G	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-5	130-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
130-6	130-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2476	134-A	·	·	·	·	·	·	·	·	A	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2488	134-A	·	·	·	·	·	·	·	·	A	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2452	134-A	·	·	·	·	·	·	·	·	A	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2493	134-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2459	134-B	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-1	134-C	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-4	134-D	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222[T])	INF	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	C	T	C	·	·	·	·	·	·	·	·	·	T	·
B. longum subsp. *longum* (JCM 1217[T])	LON	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued-7).

Strain ID	ST	Gene names *rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*	*rpoB*
		Site number[a] 81	87	105	108	123	126	136	137	150	153	165	174	189	198	219	234	240	294	309	318	324	336	369	372	390	393	402	408	414	420
31	4-A	C	T	C	T	C	T	A	G	C	T	A	T	A	G	A	T	G	T	C	T	C	C	T	C	C	C	T	C	T	A
35	4-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
1800	4-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-2	4-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-4	4-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
237	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
818	12-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
238	12-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
856	12-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
239	12-C	T	C	·	C	T	A	T	C	G	C	G	·	·	·	·	C	·	C	·	·	·	·	·	·	·	·	C	·	C	G
241	12-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
244	12-E	T	C	·	C	T	A	T	C	G	C	G	·	·	·	·	C	·	C	·	·	·	·	·	·	·	·	C	·	C	G
246	12-F	T	C	·	C	T	A	T	C	G	C	G	·	·	·	·	C	·	C	·	·	·	·	·	·	·	·	C	·	C	G
248	12-G	T	C	·	C	T	A	T	C	G	C	G	·	·	·	·	C	·	C	·	·	·	·	·	·	·	·	C	·	C	G
334	12-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
336	12-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
339	12-J	T	C	·	C	T	A	T	C	G	C	G	·	·	·	·	C	·	C	·	·	·	·	·	·	·	·	C	·	C	G
12-4	12-K	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
12-5	12-L	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·	·	·	·	·	C	G	·	T	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·	·	·	·	·	C	G	·	T	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·
139	13-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	T	·	·	·	·	C	·	·	·
141	13-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	T	·	·	·	·	C	·	·	·
144	13-D	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·
220	13-E	·	·	·	·	·	·	·	·	·	·	·	C	G	·	T	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·
1847	13-F	·	·	·	·	·	·	·	·	·	·	·	C	G	·	T	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
13-1	13-H	T	C	T	C	T	A	T	C	G	C	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	T	·	·
13-T2	13-I	·	·	·	·	·	·	·	·	·	·	·	C	G	·	T	C	·	C	·	·	·	·	·	·	·	·	·	·	·	·
281	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
401	16-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
286	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
435	16-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
297	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
441	16-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
326	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
520	16-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
298	16-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
296	16-F	T	C	·	C	T	A	T	C	G	C	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	·	·	·
485	16-G	T	C	·	C	T	A	T	C	G	C	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	·	·	·
521	16-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
527	16-I	T	C	·	C	T	A	T	C	G	C	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	·	·	·
724	16-J	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
1305	16-K	T	C	·	C	T	A	T	C	G	C	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	·	·	·
1314	16-L	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
16-1	16-M	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
16-T1	16-N	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
865	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
1120	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
1227	18-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
18-1	18-B	T	C	T	C	T	A	T	C	G	C	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	T	·	·
1521	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
30-1	30-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
630	30-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
631	30-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
1082	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	C	·	T	T	C	·	·	·
1699	38-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	C	·	T	T	C	·	·	·
1835	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1688	38-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1078	38-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	·	·	·	·	·	·
1687	38-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	·	·	·	·	·	·
38-1	38-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	C	·	·	C	·	T	T	C	·	·	·
977	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1268	44-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1426	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1854	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
44-1	44-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
1620	48-A	T	C	T	C	·	·	·	·	·	·	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	T	·	·
1841	48-A	T	C	T	C	·	·	·	·	·	·	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	T	·	·
1362	48-B	T	C	·	C	T	A	T	C	·	·	·	·	·	·	·	C	·	C	·	C	·	·	·	·	·	·	·	·	·	·
1844	48-C	T	C	T	C	·	·	·	·	·	·	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	T	·	·
48-1	48-D	T	C	T	C	·	·	·	·	·	·	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	T	·	·
48-2	48-E	T	C	T	C	·	·	·	·	·	·	G	C	G	·	T	C	A	C	·	·	·	·	·	·	·	·	·	T	·	·
2285	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
129-4	129-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2296	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
2317	130-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
2297	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
2316	130-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
2302	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
2315	130-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
2295	130-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
2308	130-E	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·
130-1	130-F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	C	·	T	·	·	·	·	C	·	·	·
130-3	130-G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	C	·	T	·	·	·	·	C	·	·	·
130-5	130-H	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	C	·	T	·	·	·	·	C	·	·	·
130-6	130-I	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	C	·	T	·	·	·	·	C	·	·	·
2476	134-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
2488	134-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
2452	134-A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	T	C	T	·	C	·	T	T	C	·	·	·
2493	134-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	C	·	T	T	C	·	·	·
2459	134-B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	C	·	·	·	·	C	·	T	T	C	·	·	·
134-1	134-C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
134-4	134-D	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222[T])	INF	·	C	·	C	T	A	T	C	G	C	G	C	G	·	·	C	A	C	·	·	·	·	·	·	·	·	·	·	·	·
B. longum subsp. *longum* (JCM 1217[T])	LON	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

Table S4. Allelic profiles for the representative strains and the type strains of *B. longum* subsp. *infantis* and *B. longum* subsp. *longum* (Continued-8).

Strain ID	ST	Gene names *rpoB* Site number[a] 423	*rpoB* 447	*rpoB* 450	*rpoB* 462	*rpoB* 471	*rpoB* 477	*rpoB* 483
31	4-A	C	C	T	C	G	T	A
35	4-B	·	·	·	·	·	C	G
1800	4-C	·	·	·	·	·	C	·
4-2	4-D	·	·	·	·	·	·	·
4-4	4-E	·	·	·	·	·	·	·
4-6	4-F	·	·	·	·	·	·	·
237	12-A	·	·	·	·	·	C	G
818	12-A	·	·	·	·	·	C	G
238	12-B	·	·	·	·	·	·	·
856	12-B	·	·	·	·	·	·	·
239	12-C	·	T	C	·	·	C	G
241	12-D	·	·	·	·	·	·	·
244	12-E	·	T	C	·	·	C	G
246	12-F	·	T	C	·	·	C	G
248	12-G	·	T	C	·	·	C	G
334	12-H	·	·	·	·	·	·	·
336	12-I	T	·	·	·	·	C	G
339	12-J	·	T	C	·	·	C	G
12-4	12-K	·	·	·	·	·	·	·
12-5	12-L	·	·	·	·	·	·	·
1784	13-A	·	·	·	·	·	·	·
13-T1	13-A	·	·	·	·	·	·	·
139	13-B	·	·	C	·	·	C	G
141	13-C	·	·	C	·	·	C	G
144	13-D	·	·	·	·	·	·	·
220	13-E	·	·	·	·	·	·	·
1847	13-F	·	·	·	·	·	·	·
1848	13-G	·	·	·	·	·	C	·
13-1	13-H	·	·	·	·	A	C	G
13-T2	13-I	·	·	·	·	·	·	·
281	16-A	T	·	·	·	·	C	G
401	16-A	T	·	·	·	·	C	G
286	16-B	T	·	·	·	·	C	G
435	16-B	T	·	·	·	·	C	G
297	16-C	T	·	·	·	·	C	G
441	16-C	T	·	·	·	·	C	G
326	16-D	T	·	·	·	·	C	G
520	16-D	T	·	·	·	·	C	G
298	16-E	T	·	·	·	·	C	G
296	16-F	·	·	·	·	·	C	G
485	16-G	·	·	·	·	·	C	G
521	16-H	T	·	·	·	·	C	G
527	16-I	·	·	·	·	·	C	G
724	16-J	T	·	·	·	·	C	G
1305	16-K	·	·	·	·	·	C	G
1314	16-L	T	·	·	·	·	C	G
16-1	16-M	T	·	·	·	·	C	G
16-T1	16-N	T	·	·	·	·	C	G
865	18-A	·	·	·	·	·	·	·
1120	18-A	·	·	·	·	·	·	·
1227	18-A	·	·	·	·	·	·	·
18-1	18-B	·	·	·	·	A	C	G
1521	30-A	·	·	·	·	·	·	·
30-1	30-A	·	·	·	·	·	·	·
630	30-B	·	·	·	·	·	C	·
631	30-C	·	·	·	·	·	C	·
1082	38-A	·	·	·	·	·	C	G
1699	38-A	·	·	·	·	·	C	G
1835	38-B	·	·	·	·	·	·	·
1688	38-B	·	·	·	·	·	·	·
1078	38-C	·	·	·	·	·	C	G
1687	38-D	·	·	·	·	·	C	G
38-1	38-E	·	·	·	·	·	C	G
977	44-A	·	·	·	·	·	·	·
1268	44-A	·	·	·	·	·	·	·
1426	44-B	·	·	·	·	·	·	·
1854	44-B	·	·	·	·	·	·	·
44-1	44-B	·	·	·	·	·	·	·
1620	48-A	·	·	·	·	A	C	G
1841	48-A	·	·	·	·	A	C	G
1362	48-B	·	·	C	·	·	C	G
1844	48-C	·	·	·	·	A	C	G
48-1	48-D	·	·	·	·	A	C	G
48-2	48-E	·	·	·	·	A	C	G
2285	129-A	·	·	·	·	·	·	·
2220	129-A	·	·	·	·	·	·	·
129-4	129-B	·	·	·	·	·	·	·
2296	130-A	·	·	·	·	·	C	·
2317	130-A	·	·	·	·	·	C	·
2297	130-B	·	·	·	·	·	·	·
2316	130-B	·	·	·	·	·	·	·
2302	130-C	·	·	·	·	·	C	·
2315	130-C	·	·	·	·	·	C	·
2295	130-D	·	·	·	·	·	·	·
2308	130-E	·	·	·	·	·	·	·
130-1	130-F	·	·	C	T	·	C	G
130-3	130-G	·	·	C	T	·	C	G
130-5	130-H	·	·	C	T	·	C	G
130-6	130-I	·	·	C	T	·	C	G
2476	134-A	·	·	·	·	·	C	G
2488	134-A	·	·	·	·	·	C	G
2452	134-A	·	·	·	·	·	C	G
2493	134-B	·	·	·	·	·	C	G
2459	134-B	·	·	·	·	·	C	G
134-1	134-C	·	·	·	·	·	C	G
134-4	134-D	·	·	·	·	·	·	·
B. longum subsp. *infantis* (JCM 1222^T)	INF	·	·	C	·	·	C	G
B. longum subsp. *longum* (JCM 1217^T)	LON	·	·	·	·	·	·	·

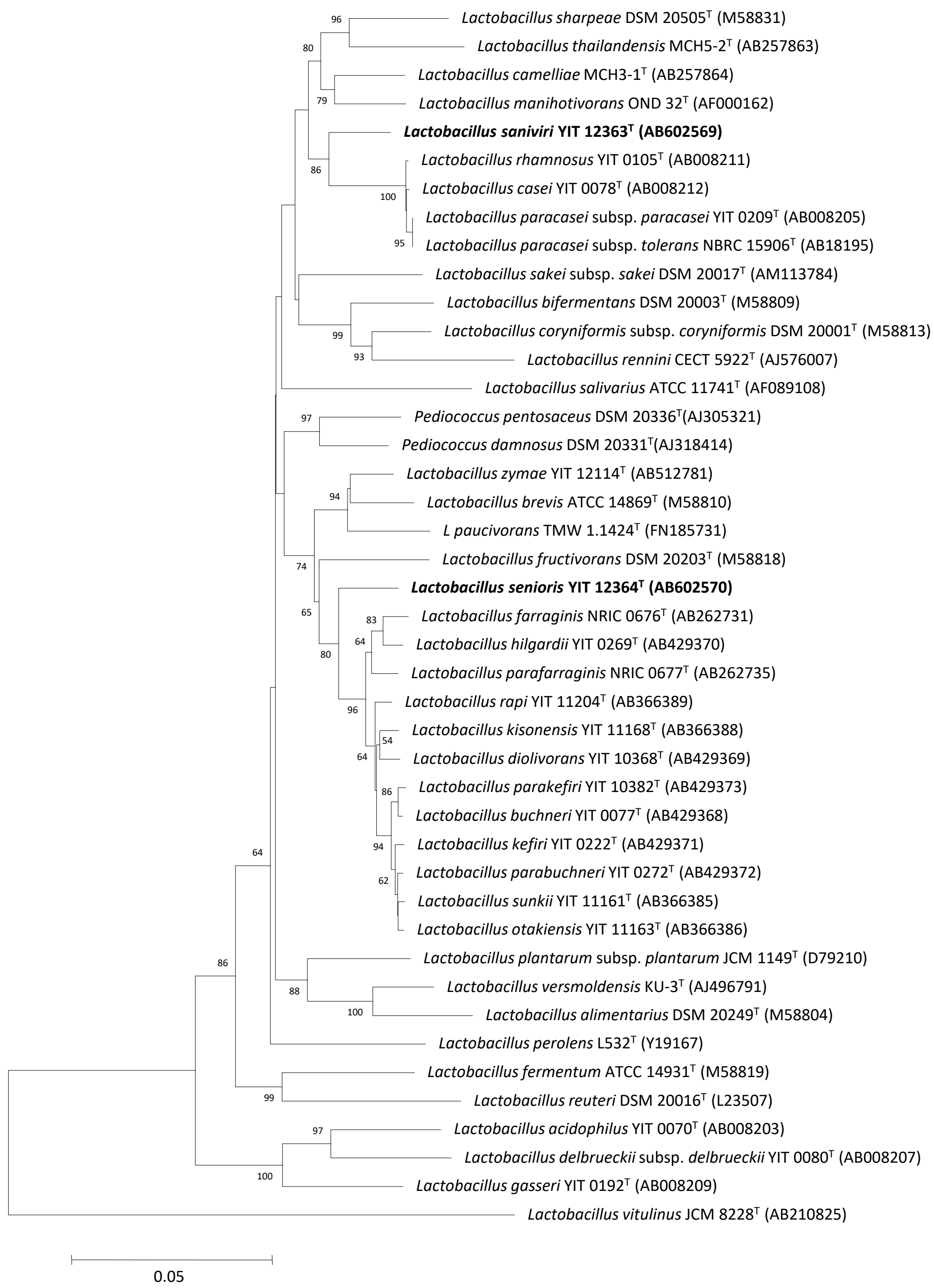


Fig. S1. Phylogenetic tree based on 16S rRNA gene sequences showing the relationship of *L. saniviri* sp. nov. YIT 12363^T and *L. senioris* sp. nov. YIT 12364^T with their closely related species.

The tree was constructed by the minimum-evolution method on the basis of a comparison of approximately 1390 bp and *L. vitulinus* JCM 8228^T was used as an outgroup. Bootstrap values (%) based on 1000 replications are given at nodes. Bar, 5% sequence divergence.

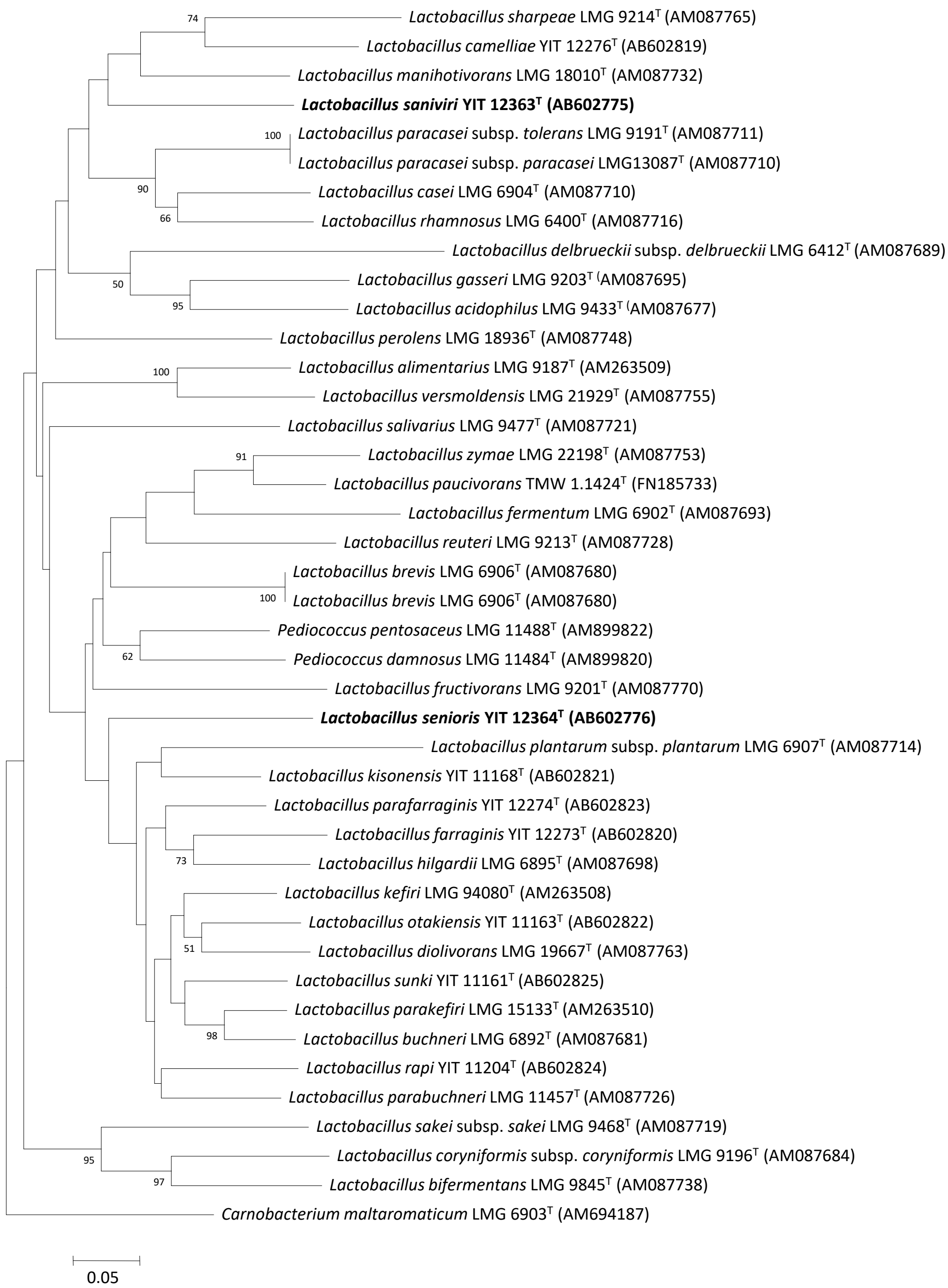


Fig. S2. Phylogenetic tree based on *pheS* gene sequences showing the relationship of *L. saniviri* sp. nov. YIT 12363^{T} and *L. senioris* sp. nov. YIT 12364^{T} with their closely related species.

The tree was constructed by the neighbor-joining method on the basis of a comparison of approximately 280 bp, and *Carnobacterium maltaromaticum* LMG 6903^{T} was used as an outgroup. Bootstrap values (%) based on 1000 replications are given at nodes. Bar, 1% sequence divergence.

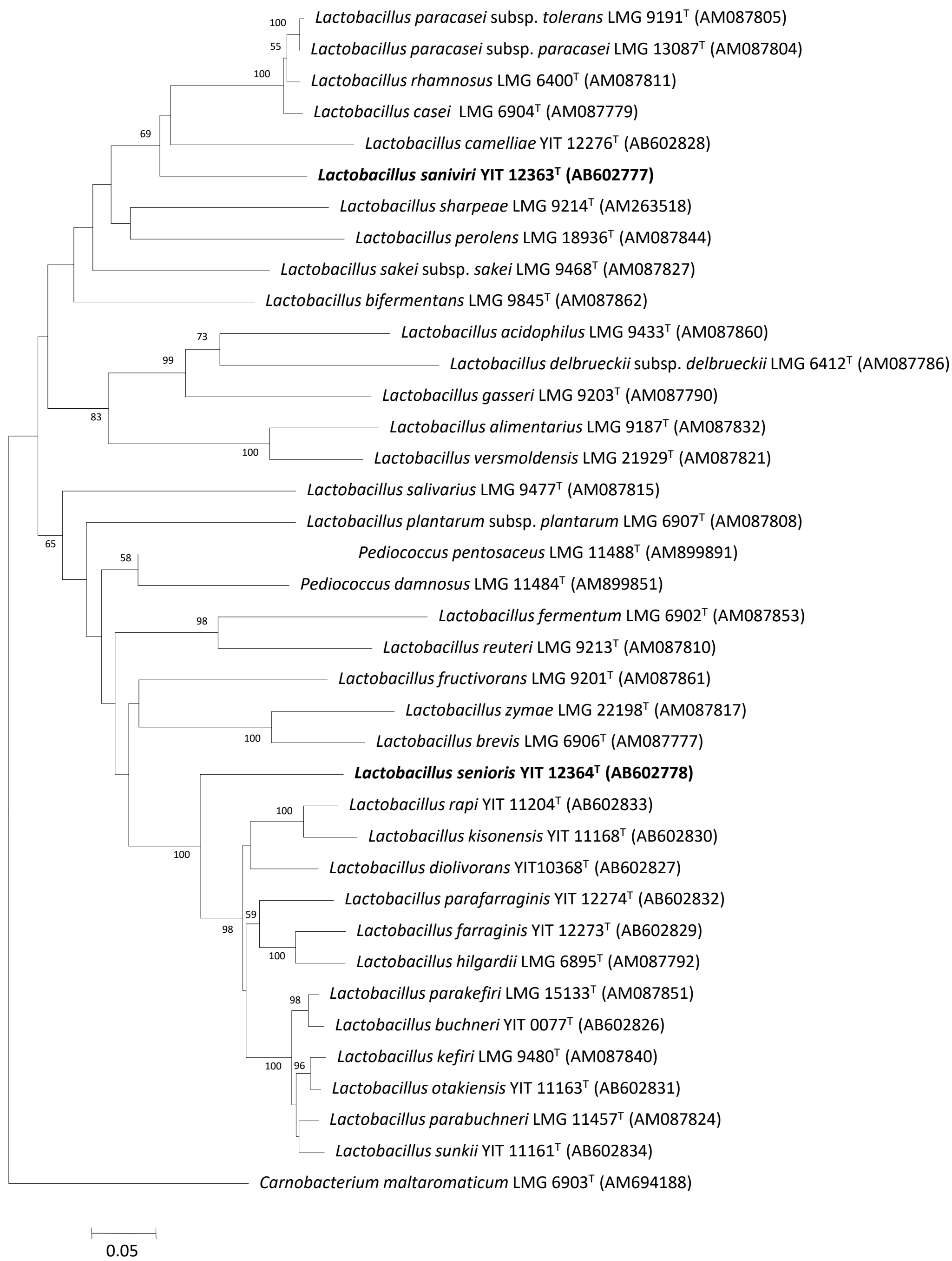


Fig. S3. Phylogenetic trees based on *rpoA* gene sequences showing the relationship of *L. saniviri* sp. nov. YIT 12363^{T} and *L. senioris* sp. nov. YIT 12364^{T} with their closely related species.

The tree was constructed by the neighbor-joining method on the basis of a comparison of approximately 350 bp, and *Carnobacterium maltaromaticum* LMG 6903^{T} was used as an outgroup. Bootstrap values (%) based on 1000 replications are given at nodes. Bar, 1% sequence divergence.

clpC

	3	12	45	111	150	163	228	234	237	240	261	266	363	366	402	423	435	468	494	501	502	504	507	528	534	537	549	550	555	558	561	564	567	570	573	582	583	600	609	624	636	637	639
clpC–1	T	C	G	C	T	G	G	G	A	C	C	C	A	C	C	C	C	T	A	A	G	C	A	C	C	G	C	C	C	C	T	C	C	T	T	T	C	C	C	C	C	G	C
clpC–2	·	·	·	T	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·
clpC–3	·	·	C	·	C	·	A	A	G	T	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	T	·	·	·	·
clpC–4	·	·	C	·	C	·	·	·	·	·	·	T	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	T	·	·	·	·
clpC–5	·	·	·	·	C	·	·	·	·	·	·	·	·	·	T	·	T	C	G	G	A	G	G	T	T	A	T	T	A	G	C	·	A	C	C	A	T	·	·	·	·	·	·
clpC–6	C	T	·	·	C	·	·	·	·	·	·	T	·	·	·	·	·	C	G	G	A	G	G	T	T	A	T	T	A	G	C	·	·	·	·	·	·	·	T	·	·	·	·
clpC–7	·	·	·	T	C	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·
clpC–8	·	·	·	·	C	·	·	·	·	·	·	·	·	·	T	T	·	C	G	G	A	G	G	·	·	·	·	·	·	·	C	·	·	·	·	·	·	T	·	T	T	A	G
clpC–9	·	·	·	·	C	C	·	·	·	·	·	T	·	T	·	·	·	C	G	G	A	G	G	T	T	A	T	T	A	·	C	·	A	C	C	A	T	·	·	·	·	·	·
clpC–10	·	·	C	·	C	·	·	·	·	·	·	T	·	·	·	·	T	·	·	·	·	·	·	T	T	A	T	T	G	G	C	T	A	C	C	A	T	·	T	·	·	·	·
clpC–11	·	·	C	·	C	·	·	·	·	·	·	T	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	G

dnaG

	15	66	72	131	142	143	144	147	159	168	171	201	204	213	223	237	238	258	261	282	285	300	306	315	321	333	357	360	363	396	429	435	441	447	465	483	492	493	495	498
dnaG–1	C	C	T	G	A	A	C	C	C	C	C	A	C	C	T	A	C	C	A	C	C	T	C	C	T	T	G	T	G	C	C	C	G	T	C	C	A	C	A	C
dnaG–2	·	·	C	·	G	C	G	·	·	·	G	G	·	·	·	·	·	T	C	T	T	C	·	G	C	C	·	·	·	T	·	A	C	C	·	·	·	·	·	·
dnaG–3	·	·	C	·	G	C	G	·	·	·	G	G	·	·	·	·	·	T	C	T	T	C	·	G	C	C	A	·	·	T	·	A	C	C	·	·	G	·	G	T
dnaG–4	T	·	C	·	·	G	·	·	T	·	G	G	·	·	C	G	·	T	C	T	T	·	·	·	·	·	C	C	·	T	T	·	·	·	T	·	·	T	·	·
dnaG–5	T	·	C	·	·	G	·	·	·	·	G	G	·	·	C	G	T	T	C	T	T	·	·	·	·	·	C	C	·	T	T	·	A	·	·	T	·	T	·	·
dnaG–6	·	·	C	·	G	C	G	·	·	·	G	G	·	·	·	·	·	T	C	T	T	C	·	G	C	C	A	·	·	T	·	A	C	C	·	·	G	·	G	T
dnaG–7	·	·	C	·	G	C	G	·	·	·	G	G	·	·	·	·	·	T	C	T	T	C	·	G	C	C	·	·	·	T	·	A	C	·	·	·	G	·	G	T
dnaG–8	·	·	C	·	G	C	G	·	·	·	G	G	T	·	·	·	·	T	C	T	T	C	·	G	C	C	A	·	·	T	·	A	C	C	·	·	·	·	·	·
dnaG–9	T	·	C	·	·	G	·	·	·	·	G	G	·	·	C	G	·	T	C	T	T	·	·	·	·	·	C	C	·	T	T	·	·	·	·	T	·	T	·	·
dnaG–10	T	T	C	A	·	G	·	T	·	·	G	G	·	·	C	G	·	·	·	T	T	·	·	·	·	·	C	C	·	T	T	·	·	C	·	T	·	T	·	·
dnaG–11	·	·	C	·	G	C	G	·	·	·	G	G	·	·	·	·	·	T	C	T	T	C	·	G	C	C	·	·	·	T	·	A	C	C	·	·	G	·	G	T
dnaG–12	T	T	C	·	·	G	·	·	·	·	G	G	·	·	C	·	·	T	C	T	T	·	T	·	·	·	C	C	A	T	T	·	·	·	·	T	·	T	·	·
dnaG–13	T	·	C	·	·	G	·	·	·	·	G	G	·	·	C	G	·	·	·	T	T	C	·	·	C	·	C	C	·	T	·	·	C	C	·	T	·	T	·	·
dnaG–14	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
dnaG–15	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	G	T
dnaG–16	T	·	C	·	·	G	·	·	T	·	G	G	·	·	C	G	·	T	C	T	T	·	·	·	·	·	C	C	·	T	T	·	·	·	T	·	·	T	·	·
dnaG–17	·	·	C	·	·	·	·	·	·	·	·	·	·	T	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
dnaG–18	·	·	C	·	·	·	·	·	·	T	·	·	·	T	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·

	501	507	540	564	588	612	618	661	669	679	693	723	756	759	774	778	792	798	810	813	828	831	834	837	843	846	849	861	864	873
dnaG–1	A	T	C	G	G	A	G	G	C	A	C	T	C	G	G	T	C	G	C	C	T	T	C	C	C	G	A	T	C	C
dnaG–2	·	·	·	·	C	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T
dnaG–3	G	C	T	·	C	G	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	T	·
dnaG–4	·	C	T	·	·	G	C	·	·	·	·	·	·	·	T	C	T	A	T	T	C	·	·	T	T	·	G	·	·	T
dnaG–5	·	C	·	A	·	G	C	T	·	·	·	·	·	·	·	C	T	A	T	T	C	·	·	T	T	·	G	·	·	T
dnaG–6	G	C	T	·	C	G	·	·	·	·	·	·	T	A	·	·	·	·	·	·	·	·	·	·	·	·	·	C	T	·
dnaG–7	G	C	T	·	C	G	·	·	·	G	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	T	·
dnaG–8	·	·	·	·	C	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T
dnaG–9	·	C	·	·	·	G	C	·	·	·	·	·	·	·	·	C	·	A	T	T	C	·	·	T	T	T	G	C	·	T
dnaG–10	·	C	·	·	·	G	C	·	·	·	·	C	·	·	·	C	·	A	T	T	C	·	·	·	·	·	G	·	·	T
dnaG–11	G	C	T	·	C	G	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	T	·
dnaG–12	·	C	·	·	·	G	C	·	·	·	·	·	·	·	·	C	·	A	T	T	C	·	·	T	T	T	G	C	·	T
dnaG–13	·	C	·	·	·	G	C	·	·	·	·	C	·	·	·	C	·	A	T	T	C	A	·	·	·	·	G	·	·	T
dnaG–14	·	·	T	·	·	G	·	·	T	·	T	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T
dnaG–15	G	C	·	·	·	·	·	·	T	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T
dnaG–16	·	C	T	·	·	G	C	·	·	·	·	C	·	·	·	C	·	A	T	T	C	·	T	T	·	·	G	·	·	T
dnaG–17	·	·	·	·	·	·	·	·	T	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T
dnaG–18	·	·	·	·	·	·	·	·	T	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T

dnaJ1

	8	26	32	38	65	80	83	89	92	101	116	122	125	131	158	217	224	234	239	240	241	243	282	329
dnaJ1−1	T	T	C	C	T	C	C	T	T	C	C	T	C	C	C	A	T	G	A	T	C	C	C	C
dnaJ1−2	·	C	A	·	·	·	·	·	·	·	·	·	·	·	T	·	A	A	G	C	A	G	·	·
dnaJ1−3	·	C	A	·	·	·	·	·	·	·	·	·	·	·	·	·	A	A	G	C	A	G	·	·
dnaJ1−4	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
dnaJ1−5	·	C	·	·	·	·	·	C	·	·	·	·	·	·	·	·	A	A	G	C	A	G	·	·
dnaJ1−6	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	A	G	C	A	G	·	·
dnaJ1−7	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·
dnaJ1−8	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	•	G	C	A	G	T	·
dnaJ1−9	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	A	G	C	A	G	·	T
dnaJ1−10	C	C	·	·	·	·	·	·	·	·	·	C	T	·	·	·	A	A	G	C	A	G	·	·
dnaJ1−11	C	·	·	A	C	T	T	C	C	T	T	C	T	T	·	G	·	·	·	·	·	·	·	·

hsp60

	11	50	53	65	89	92	113	143	173	236	245	302	312	335	416
hsp60–1	C	C	T	C	T	T	T	G	T	T	C	T	C	A	C
hsp60–2	·	·	·	·	·	·	C	·	·	C	·	·	·	C	T
hsp60–3	·	·	·	·	·	·	C	·	·	C	G	·	·	C	T
hsp60–4	·	·	·	·	·	·	·	·	·	C	·	C	T	C	·
hsp60–5	·	·	·	·	C	·	C	·	·	C	·	·	·	C	T
hsp60–6	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·
hsp60–7	T	·	·	·	·	·	C	·	·	C	·	·	·	C	T
hsp60–8	·	·	·	·	C	C	C	·	·	C	·	·	·	C	T
hsp60–9	·	T	C	T	·	·	·	·	C	C	·	·	·	C	·
hsp60–10	·	·	·	·	·	·	C	·	·	C	·	·	·	C	·

Fig. S4. Polymorphic nucleotide sites found among the 25 *B. longum* studied strains at the seven MLST gene loci.

For each gene, all discovered alleles were compared, but only polymorphic sites are shown. Numbering starts at the beginning of the aligned sequence portion of each gene.

purF

	14	26	32	38	53	62	68	80	86	101	119	120	122	134	140	152	161	164	167	170	176	182	188	191	197	203	206	209	219	227	238	258	287	302	305	317	323	326	327	332	338	353
purF–1	A	G	T	T	G	A	C	A	C	C	G	A	G	C	G	A	A	G	T	C	C	G	G	T	T	G	T	A	G	T	A	A	G	C	C	T	T	C	G	C	C	C
purF–2	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–3	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·
purF–4	·	C	C	C	C	G	G	G	·	T	·	·	·	·	A	·	G	C	C	G	·	A	A	C	C	C	·	G	·	C	·	·	A	·	T	C	C	·	·	·	·	·
purF–5	·	C	C	C	C	G	G	G	T	T	·	·	·	·	A	G	C	C	C	G	T	·	A	C	C	C	C	·	·	C	·	·	A	·	·	C	C	T	·	·	T	T
purF–6	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·
purF–7	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	C	C	G	·	A	·	C	C	C	C	·	·	C	·	·	A	·	T	C	C	·	·	·	·	·
purF–8	·	C	C	C	C	G	G	G	T	T	·	·	·	·	·	G	C	C	C	G	·	A	·	C	·	C	C	·	·	C	G	·	A	·	T	C	C	·	·	·	·	·
purF–9	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	C	·	·	·	·	·	A	A	·	·	·	T	·	·	·	·
purF–10	G	C	C	C	C	G	G	G	T	T	·	·	·	·	·	G	C	C	C	G	T	A	A	C	C	C	C	·	·	·	·	·	A	A	·	·	·	·	·	·	·	·
purF–11	·	C	C	C	C	G	G	G	T	T	·	·	·	·	·	G	C	C	C	G	·	A	·	C	·	C	C	·	·	C	G	·	A	·	T	C	C	·	·	·	·	·
purF–12	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–13	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	C	·	·	·	·	T	A	A	·	·	·	·	·	·	·	·
purF–14	·	C	C	C	C	G	G	G	T	T	A	C	C	A	·	G	C	C	C	G	·	·	A	C	C	C	C	·	A	C	·	·	A	·	T	C	C	·	·	·	T	·
purF–15	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	C	·	·	·	·	·	A	A	·	·	·	·	·	·	·	·
purF–16	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	C	·	·	·	·	·	A	A	·	·	·	·	·	T	·	·
purF–17	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	C	·	·	·	·	·	A	A	·	·	·	·	·	·	·	·

	359	362	368	371	372	373	374	377	386	389	395	401	422	423	431	434	440	443	449	452	455	458	470	476	479	485	494	503	506	512	521	524	528	536	542	545	548	557	560	563	569	570
purF–1	C	A	C	C	G	C	C	A	G	G	T	T	A	T	C	A	C	T	A	A	A	C	A	T	A	A	T	C	C	T	G	A	C	G	T	G	T	C	T	T	C	A
purF–2	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–3	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–4	·	G	G	G	C	A	G	G	C	·	·	C	C	C	G	G	·	C	G	C	C	G	G	C	G	G	C	T	G	C	A	G	·	A	C	·	C	G	C	C	T	G
purF–5	T	G	G	G	C	A	G	G	C	T	C	C	C	C	G	G	·	C	G	C	T	A	G	·	G	G	C	·	·	C	A	G	·	A	A	T	C	G	C	C	T	G
purF–6	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–7	·	G	G	G	C	A	G	G	C	·	·	C	C	C	G	G	·	C	G	C	C	G	·	·	G	G	C	·	·	C	A	G	·	A	C	·	C	G	C	C	T	G
purF–8	·	G	G	G	C	A	G	G	C	·	·	C	C	C	G	C	·	C	·	G	T	G	G	·	G	G	C	T	G	C	A	G	T	·	C	·	C	G	C	C	·	G
purF–9	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·
purF–10	·	·	G	G	C	A	G	G	C	·	·	C	C	C	G	G	·	C	G	C	·	G	·	C	G	G	C	T	G	C	A	G	T	·	C	·	C	G	C	C	·	G
purF–11	·	G	G	G	C	A	G	G	C	·	·	C	C	C	·	·	T	·	G	C	C	G	G	C	G	G	C	T	G	C	A	G	T	·	C	·	C	G	C	C	·	G
purF–12	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–13	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·
purF–14	·	G	G	G	C	A	G	G	C	T	C	C	C	C	G	G	·	C	G	C	C	G	G	C	G	G	C	T	G	C	A	G	T	·	C	·	C	G	C	C	·	G
purF–15	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·
purF–16	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·
purF–17	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·

	578	581	590	593	594	599	605	611	620	629	644	648	650	653	654	655	656	659	665	671	680	686	695	701	704	707	713	722	728	731	749	758
purF–1	T	A	C	T	A	C	T	C	C	G	G	C	G	C	A	A	G	A	C	T	T	C	C	G	G	G	C	T	C	T	A	A
purF–2	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–3	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–4	C	G	·	G	G	·	C	G	T	A	·	·	·	G	C	G	C	·	·	C	C	T	G	·	C	·	G	C	·	C	G	C
purF–5	C	G	·	G	G	·	C	G	T	A	·	·	·	G	C	G	C	·	·	C	C	T	G	·	C	·	G	C	·	C	G	C
purF–6	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–7	C	·	·	G	G	·	C	G	T	C	·	·	·	G	C	G	C	·	·	C	C	·	G	A	C	A	G	C	·	C	·	C
purF–8	C	·	·	G	G	·	C	G	·	C	C	T	·	G	C	G	C	G	·	C	C	·	G	·	C	·	G	C	·	C	G	C
purF–9	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	T	·	·	·
purF–10	C	·	G	G	G	·	C	G	·	·	·	·	A	G	C	G	C	·	·	C	C	·	A	·	C	A	G	C	·	C	·	C
purF–11	C	·	G	G	G	·	C	G	·	C	C	T	·	G	C	G	C	G	T	C	C	T	G	·	C	A	G	C	·	C	·	C
purF–12	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
purF–13	·	·	·	·	·	T	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	T	·	·	·
purF–14	C	·	G	G	G	·	C	G	·	·	·	·	A	G	C	G	C	·	·	C	C	·	G	·	C	A	G	C	·	C	·	C
purF–15	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	T	·	T	·	·	·
purF–16	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	T	·	·	·
purF–17	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	T	·	·	·

rpoC

	63	99	114	126	141	189	195	255	300	318	342	369	396	456	468	495	508	539	552	558	559	560	573	579	585	633	642	645	651	654	765	768	798	843	867	894	897	903	909	912	969	1050	1062	1089
rpoC–1	G	G	G	G	A	A	G	G	G	T	G	T	C	G	G	G	A	A	A	G	C	T	G	G	G	G	T	G	G	C	A	A	G	C	A	C	A	C	C	G	G	A	C	G
rpoC–2	·	·	·	·	·	·	·	·	·	·	A	C	·	·	·	A	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	T	·	·	·	·	·	·	·	·	·	-
rpoC–3	A	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	-
rpoC–4	·	·	A	A	·	G	·	A	·	C	A	C	T	·	·	·	G	G	G	·	G	A	·	A	A	·	·	·	·	·	·	·	·	·	G	T	G	·	·	·	A	·	·	-
rpoC–5	·	·	A	A	G	·	A	A	·	C	A	C	T	·	·	·	G	G	G	·	G	A	·	A	A	·	·	·	·	·	·	·	·	·	G	T	G	·	·	·	A	·	·	A
rpoC–6	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	-
rpoC–7	·	·	A	A	·	·	·	·	·	C	A	C	T	·	·	·	G	G	G	·	G	A	·	A	A	·	C	·	·	G	·	·	·	·	G	·	G	·	·	·	·	·	·	-
rpoC–8	·	·	·	A	·	·	A	A	·	C	A	C	T	·	·	·	G	G	G	·	G	A	·	A	A	·	·	·	C	G	·	·	·	·	G	·	G	·	·	·	A	·	T	-
rpoC–9	·	A	·	A	·	·	A	·	·	C	A	C	A	·	·	·	G	G	G	·	G	A	·	A	A	·	C	·	C	G	·	·	·	·	G	T	G	G	A	C	A	·	·	-
rpoC–10	·	A	·	A	·	·	A	·	·	C	A	C	A	·	·	·	G	G	G	·	G	A	·	A	A	A	·	·	·	·	·	·	·	·	G	·	G	·	·	·	·	·	·	-
rpoC–11	·	·	A	A	·	·	A	A	·	C	A	C	T	·	·	·	G	G	G	·	G	A	·	A	A	·	·	·	·	·	G	G	·	·	G	·	G	·	·	·	A	G	·	-
rpoC–12	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	-
rpoC–13	·	·	·	A	·	·	·	A	·	C	A	C	A	·	·	·	G	G	G	·	G	A	·	A	A	·	C	·	C	G	·	·	·	·	G	·	G	·	·	·	A	·	T	-
rpoC–14	·	·	A	A	·	G	·	·	·	C	A	·	A	A	·	·	G	G	G	·	·	·	·	A	A	·	C	·	·	G	·	·	·	T	G	·	G	·	·	·	·	·	·	-
rpoC–15	·	·	A	A	·	·	·	A	·	C	A	C	A	·	A	·	G	G	G	A	G	A	·	A	A	·	·	·	C	G	·	·	·	·	G	·	G	·	·	·	·	·	·	-

xfp

	48	64	67	108	109	112	116	124
xfp–1	A	C	T	C	C	C	C	T
xfp–2	·	·	C	·	·	·	·	·
xfp–3	·	·	C	·	T	·	·	·
xfp–4	G	·	C	·	·	·	·	C
xfp–5	G	·	C	·	·	T	·	C
xfp–6	G	T	C	·	·	·	T	C
xfp–7	·	·	·	G	·	·	·	·

Fig. S4. Polymorphic nucleotide sites found among the 25 *B. longum* studied strains at the seven MLST gene loci. (Continued)

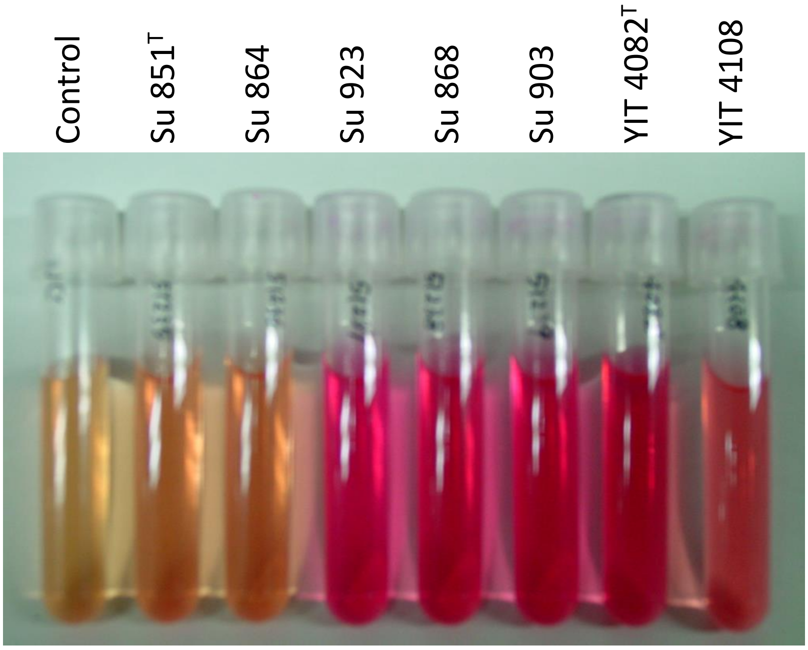


Fig. S5. Urease activity of the *B. longum* subsp. *suis* strains.

The alkaline production in 4 ml of broth was determined by the color change in the medium to red.

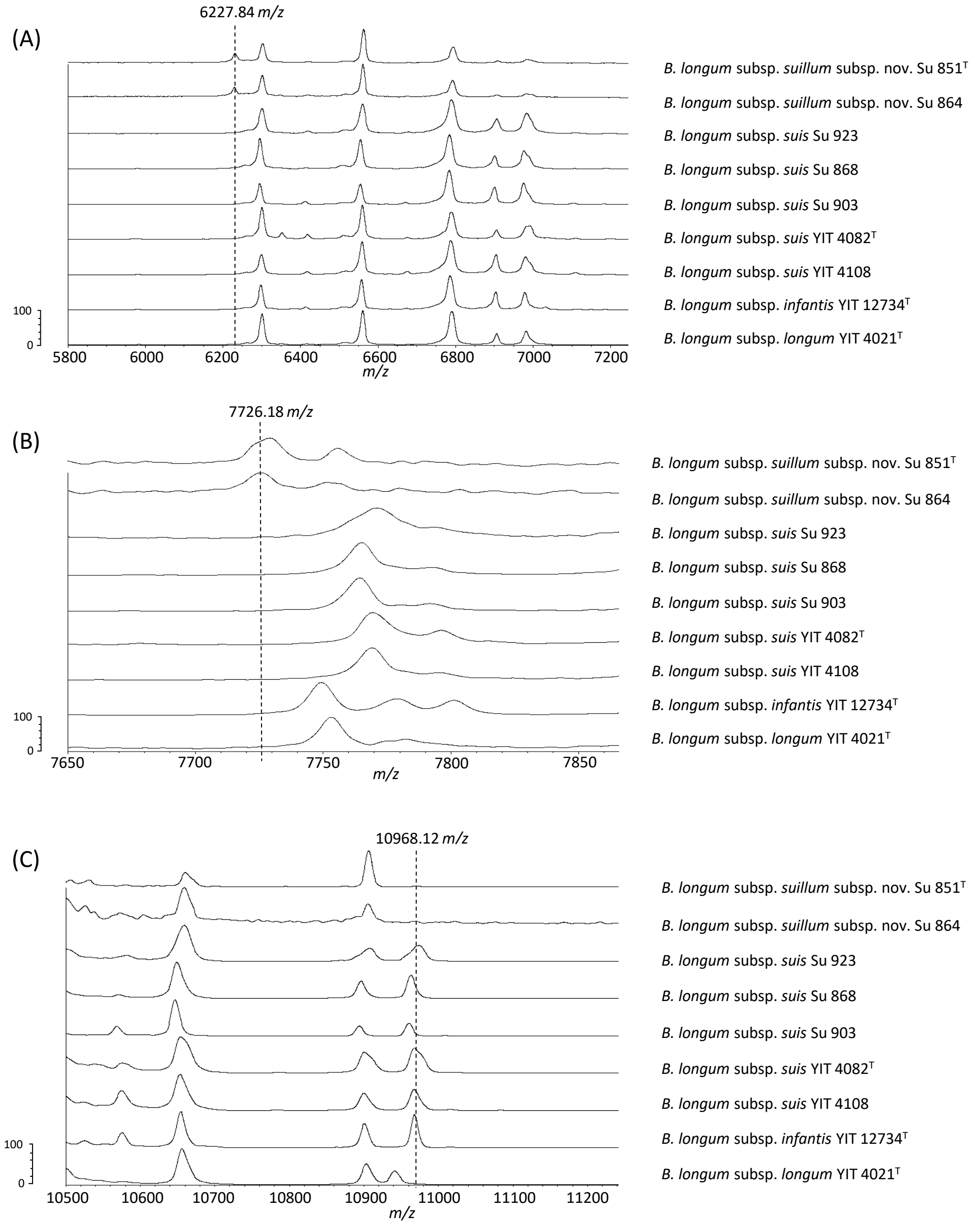


Fig. S6. MALDI-TOF MS spectra of *B. longum* strains.

MALDI-TOF MS spectra of cell lysate were shown for each strain in the range of 5800–7200 (A), 7650–7850 (B) and 10500–11200 *m/z* (C). Specific peaks for the urease-negative strains *B. longum* subsp. *suillum* subsp. nov. (Su 851^{T} and Su 864) were observed at around 6228 and 7730 m/z, respectively, but not observed at around 10970 *m/z*, whereas the urease-positive stains (five strains of *B. longum* subsp. *suis* Su 868, Su 903, Su 923, YIT 4082^{T}, YIT 4108 and type strains of *B. longum* subsp. *infantis* (YIT 12734^{T}) and *B. longum* subsp. *longum* (YIT 4021^{T}) showed contrary features. Dotted lines indicate the size of 6227.84 (A), 7726.18 (B) and 10968.12 m/z (C), respectively.

Subject 4

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (133 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	−	−	6.65	9.23	9.27	8.94	9.99	9.13
B. adolescentis	−	−	−	−	−	−	−	−
B. animalis subsp. *lactis*	−	−	6.71	9.41	9.53	−	7.63	−
B. bifidum	−	−	−	−	6.17	6.11	−	9.04
B. breve	−	−	−	−	−	−	8.81	−
B. catenulatum group	−	−	−	−	−	−	8.87	8.42
B. dentium	−	−	6.09	−	−	−	−	−
B. longum subsp. *infantis*	−	−	−	−	−	−	8.48	−
B. longum subsp. *longum*	−	−	−	−	−	9.60	8.30	8.61

Subject 12

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (> 180 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	−	−	6.65	9.23	9.27	8.94	9.99	9.13
B. adolescentis	−	−	−	−	−	−	−	−
B. animalis subsp. *lactis*	−	−	6.71	9.41	9.53	−	7.63	−
B. bifidum	−	−	−	−	6.17	6.11	−	9.04
B. breve	−	−	−	−	−	−	8.81	−
B. catenulatum group	−	−	−	−	−	−	8.87	8.42
B. dentium	−	−	6.09	−	−	−	−	−
B. longum subsp. *infantis*	−	−	−	−	−	−	8.48	−
B. longum subsp. *longum*	−	−	−	−	−	9.60	8.30	8.61

Subject 13[c]

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (120 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	8.96	10.25	10.69	10.49	9.84	9.71	10.07	9.84
B. adolescentis	−	−	−	−	6.42	−	−	10.04
B. animalis subsp. *lactis*	−	−	−	7.64	−	−	−	−
B. bifidum	−	−	−	−	9.78	10.12	9.05	−
B. breve	8.81	10.56	10.53	10.30	9.39	9.72	8.50	−
B. catenulatum group	−	−	−	−	−	−	−	−
B. dentium	−	−	−	−	−	−	−	−
B. longum subsp. *infantis*	−	−	−	−	−	−	−	−
B. longum subsp. *longum*	−	−	−	−	8.08	8.42	9.65	9.17

Subject 16

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (> 180 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	−	9.26	8.87	9.12	8.81	ns	ns	9.50
B. adolescentis	−	8.38	−	−	6.21	ns	ns	7.98
B. animalis subsp. *lactis*	−	−	−	−	−	ns	ns	−
B. bifidum	−	8.67	9.20	9.57	9.32	ns	ns	9.57
B. breve	−	−	−	−	−	ns	ns	−
B. catenulatum group	−	−	−	−	−	ns	ns	8.86
B. dentium	−	7.16	−	−	−	ns	ns	−
B. longum subsp. *infantis*	−	−	−	−	−	ns	ns	−
B. longum subsp. *longum*	−	8.55	8.33	8.67	8.50	ns	ns	8.35

Subject 18

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (136 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	−	9.64	9.22	9.22	10.24	10.20	9.48	10.00
B. adolescentis	−	−	−	6.14	−	−	−	8.16
B. animalis subsp. *lactis*	−	−	−	−	−	−	−	−
B. bifidum	−	−	−	−	−	−	9.20	−
B. breve	−	8.94	8.51	8.45	9.68	9.49	7.60	−
B. catenulatum group	−	−	−	−	−	−	−	9.57
B. dentium	−	−	−	−	−	−	9.02	−
B. longum subsp. *infantis*	−	−	−	−	−	−	−	−
B. longum subsp. *longum*	−	9.49	9.02	8.98	10.04	9.93	9.42	8.83

Subject 30[c]

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (130 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	6.11	ns	−	−	8.99	10.89	8.92	9.61
B. adolescentis	−	ns	−	−	−	7.98	−	6.18
B. animalis subsp. *lactis*	−	ns	−	−	−	7.40	6.92	−
B. bifidum	6.22	ns	−	−	9.83	9.65	7.84	−
B. breve	−	ns	−	−	−	−	8.89	−
B. catenulatum group	−	ns	−	−	8.14	10.29	5.92	9.08
B. dentium	−	ns	−	−	−	6.22	−	−
B. longum subsp. *infantis*	−	ns	−	−	−	−	−	−
B. longum subsp. *longum*	−	ns	−	−	8.70	8.74	7.07	8.39

Fig. S7. Raw data of qPCR analysis.

The abundance of genus *Bifidobacterium* as well as seven species and three subspecies of *Bifidobacterium* were measured at each time point for 12 subject whose fecal samples at least one *B. longum* subsp. *longum* strain was detected in early infancy or childhood.

[a] ns, no sample was collected; −, not detected. [b] One week after the introduction of solids.

[c] Subject from whose fecal samples at least one long-term colonizer was detected.

Subject 38

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (99 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	ns	9.42	8.99	8.59	9.66	ns	10.15	9.69
B. adolescentis	ns	8.97	6.60	8.08	8.84	ns	10.25	8.80
B. animalis subsp. *lactis*	ns	—	—	—	—	ns	—	—
B. bifidum	ns	9.92	9.49	9.74	9.25	ns	8.35	9.06
B. breve	ns	—	—	—	9.75	ns	9.53	—
B. catenulatum group	ns	—	—	—	—	ns	—	8.74
B. dentium	ns	—	—	—	—	ns	6.19	—
B. longum subsp. *infantis*	ns	—	—	—	—	ns	—	—
B. longum subsp. *longum*	ns	8.83	8.74	8.37	9.08	ns	8.87	8.54

Subject 44[c]

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (110 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	—	ns	9.65	9.04	9.70	9.54	9.78	9.73
B. adolescentis	—	ns	—	6.22	—	—	9.19	8.64
B. animalis subsp. *lactis*	—	ns	—	—	—	—	—	—
B. bifidum	—	ns	—	9.77	10.00	10.25	9.40	9.69
B. breve	—	ns	6.45	—	—	—	—	—
B. catenulatum group	—	ns	—	—	8.87	9.11	8.89	9.18
B. dentium	—	ns	6.99	—	6.06	—	—	—
B. longum subsp. *infantis*	—	ns	—	—	—	—	—	—
B. longum subsp. *longum*	—	ns	9.43	8.59	9.25	8.86	8.59	8.08

Subject 48

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (124 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	7.32	ns	8.19	9.81	9.74	9.85	9.72	10.07
B. adolescentis	—	ns	—	—	—	—	—	9.90
B. animalis subsp. *lactis*	8.18	ns	8.32	—	—	—	—	—
B. bifidum	—	ns	—	—	9.47	9.37	9.61	9.55
B. breve	—	ns	—	—	—	—	—	—
B. catenulatum group	—	ns	—	—	—	—	—	8.69
B. dentium	—	ns	—	6.46	—	—	—	—
B. longum subsp. *infantis*	—	ns	—	—	—	—	—	—
B. longum subsp. *longum*	—	ns	—	9.73	9.77	9.72	9.58	9.10

Subject 129

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (128 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	—	9.13	10.07	10.19	8.52	10.31	9.65	9.91
B. adolescentis	—	—	—	—	—	—	—	7.79
B. animalis subsp. *lactis*	—	—	—	—	—	—	6.05	—
B. bifidum	—	6.17	—	—	—	—	—	—
B. breve	—	—	—	—	—	—	—	—
B. catenulatum group	—	—	—	—	—	—	—	9.34
B. dentium	—	—	—	—	—	—	—	—
B. longum subsp. *infantis*	—	—	—	—	—	—	—	—
B. longum subsp. *longum*	—	8.92	9.95	10.16	8.92	10.16	9.63	8.84

Subject 130

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[b] (124 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	—	8.67	10.27	10.14	10.36	10.10	10.06	10.13
B. adolescentis	—	—	—	—	—	—	—	—
B. animalis subsp. *lactis*	—	—	—	—	—	—	6.40	—
B. bifidum	—	8.98	8.01	9.15	9.50	9.09	7.93	9.86
B. breve	—	8.34	10.25	10.35	10.55	10.13	9.10	6.07
B. catenulatum group	—	7.86	7.34	7.59	8.14	8.50	9.35	9.71
B. dentium	—	—	—	—	—	—	—	—
B. longum subsp. *infantis*	—	—	—	—	—	—	—	—
B. longum subsp. *longum*	—	8.08	7.53	7.73	8.16	8.63	—	9.43

Subject 134

Target	Bacterial counts in each time point (log10 cells / g feces)[a]							
	0 day	3 days	7 days	30 days	90 days	After weaning[c] (162 days)	180 days	≈ 6 years
Genus *Bifidobacterium*	—	7.97	9.06	8.77	9.51	10.59	10.46	9.68
B. adolescentis	—	6.38	—	6.16	—	—	9.13	7.05
B. animalis subsp. *lactis*	—	—	—	—	—	—	—	—
B. bifidum	—	8.72	10.03	9.48	9.58	9.98	9.71	7.61
B. breve	—	—	—	—	—	10.47	10.48	—
B. catenulatum group	—	—	—	7.56	8.78	9.41	8.86	9.12
B. dentium	—	—	—	—	—	—	—	—
B. longum subsp. *infantis*	—	—	—	—	—	—	—	—
B. longum subsp. *longum*	—	7.52	8.44	8.91	9.18	9.18	8.63	8.28

Fig. S7 Raw data of qPCR analysis (Continued).

References

1. Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, Tobe T, Clarke JM, Topping DL, Suzuki T, Taylor TD, Itoh K, Kikuchi J, Morita H, Hattori M, Ohno H. 2011. Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature **469**:543–547.
2. Makino H, Kushiro A, Ishikawa E, Kubota H, Gawad A, Sakai T, Oishi K, Martin R, Ben-Amor K, Knol J, Tanaka R. 2013. Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota. PLoS One **8**:e78331.

Notes

This thesis also refers the following articles:

Oki K, Kudo Y, Watanabe K. 2012. *Lactobacillus saniviri* sp. nov. and *Lactobacillus senioris* sp. nov., isolated from human faeces. Int J Syst Evol Microbiol **62**:601–607.

Yanokura E, Oki K, Makino H, Modesto M, Pot B, Mattarelli P, Biavati B, Watanabe K. 2015. Subspeciation of *Bifidobacterium longum* by multilocus approaches and amplified fragment length polymorphism: Description of *B. longum* subsp. *suillum* subsp. nov., isolated from the faeces of piglets. Syst Appl Microbiol **38**:305–314.

Oki K, Akiyama T, Matsuda K, Gawad A, Makino H, Ishikawa E, Oishi K, Kushiro A, Fujimoto J. 2018. Long-term colonization exceeding six years from early infancy of *Bifidobacterium longum* subsp. *longum* in human gut. BMC Microbiol **18**:209.