DNA Research **15**, 151–161, (2008) doi:10.1093/dnares/dsn009

Comparative Genome Analysis of *Lactobacillus reuteri* and *Lactobacillus fermentum* Reveal a Genomic Island for Reuterin and Cobalamin Production

Hidetoshi Morita[1,*], Hidehiro Toh[2,3], Shinji Fukuda[4,5], Hiroshi Horikawa[1], Kenshiro Oshima[3], Takehito Suzuki[1], Masaru Murakami[1], Shin Hisamatsu[6], Yukio Kato[1], Tatsuya Takizawa[1], Hideo Fukuoka[1], Tetsuhiko Yoshimura[7], Kikuji Itoh[8], Daniel J. O'Sullivan[9], Larry L. McKay[9], Hiroshi Ohno[4,5], Jun Kikuchi[5,10,11], Toshio Masaoka[1], and Masahira Hattori[2,12]

School of Veterinary Medicine, Azabu University, 1-17-71 Fuchinobe, Sagamihara, Kanagawa 229-8501, Japan[1]; RIKEN Genomic Sciences Center, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan[2]; Kitasato Institute for Life Sciences, Kitasato University, 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan[3]; RIKEN Research Center for Allergy and Immunology, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan[4]; International Graduate School of Arts and Sciences, Yokohama City University, 1-7-29 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan[5]; College of Environmental Health, Azabu University, 1-17-71 Fuchinobe, Sagamihara, Kanagawa 229-8501, Japan[6]; Yamagata Promotional Organization for Industrial Technology, 2-2-1 Matsuei, Yamagata 990-2473, Japan[7]; Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113-8657, Japan[8]; Department of Food Science and Nutrition, University of Minnesota, 1334 Eckles Avenue, St. Paul 55108, USA[9]; RIKEN Plant Science Center, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan[10]; Graduate School of Bioagriculture Sciences, Nagoya University, 1 Furo, Chikusa, Nagoya, Aichi 464-8601, Japan[11] and Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan[12]

(Received 12 March 2008; accepted on 10 April 2008; published online 16 May 2008)

Abstract

***Lactobacillus reuteri* is a heterofermentative lactic acid bacterium that naturally inhabits the gut of humans and other animals. The probiotic effects of *L. reuteri* have been proposed to be largely associated with the production of the broad-spectrum antimicrobial compound reuterin during anaerobic metabolism of glycerol. We determined the complete genome sequences of the reuterin-producing *L. reuteri* JCM 1112^T and its closely related species *Lactobacillus fermentum* IFO 3956. Both are in the same phylogenetic group within the genus *Lactobacillus*. Comparative genome analysis revealed that *L. reuteri* JCM 1112^T has a unique cluster of 58 genes for the biosynthesis of reuterin and cobalamin (vitamin B_{12}). The 58-gene cluster has a lower GC content and is apparently inserted into the conserved region, suggesting that the cluster represents a genomic island acquired from an anomalous source. Two-dimensional nuclear magnetic resonance (2D-NMR) with $^{13}C_3$-glycerol demonstrated that *L. reuteri* JCM 1112^T could convert glycerol to reuterin *in vivo*, substantiating the potential of *L. reuteri* JCM 1112^T to produce reuterin in the intestine. Given that glycerol is shown to be naturally present in feces, the acquired ability to produce reuterin and cobalamin is an adaptive evolutionary response that likely contributes to the probiotic properties of *L. reuteri*.**

Key words: *Lactobacillus reuteri*; *Lactobacillus fermentum*; reuterin; cobalamin; genome

Edited by: Katsumi Isono

* To whom correspondence should be addressed. Tel./fax: + 81 42-850-2450. E-mail: morita@azabu-u.ac.jp

1. Introduction

Lactobacillus reuteri is a heterofermentative lactic acid bacterium and is frequently found in the

gastrointestinal tract of humans and other animals. According to the World Health Organization (WHO), probiotics are defined as live microorganisms which when administered in adequate amounts confer a health benefit on the host. *Lactobacillus reuteri* has been reported to exhibit "probiotic" properties.[1,2] It has been extensively analyzed for probiotic applications, including its safe administration to healthy individuals, its ability to colonize the intestine, as a diarrhea therapeutic agent, as an inhibitor of bacterial pathogens, and the immunological modulation of the gastrointestinal mucosa.[3–8] Some strains of *L. reuteri* have the ability to produce and excrete the broad-spectrum antimicrobial compound reuterin, which is structurally identical to 3-hydroxypropionaldehyde (3-HPA), during anaerobic metabolism of glycerol.[9–11] The probiotic effects of *L. reuteri* have been proposed to be largely associated with the production of reuterin, and this antimicrobial substance is also an effective food preservative agent.[12,13] This attribute of reuterin-producing *L. reuteri* strains is why they are frequently used as probiotic cultures in commercial dairy products worldwide.[13]

The molecular basis for the probiotic effects of *L. reuteri* is not yet established. Although reuterin production is proposed to be important for its probiotic effects, the *in vivo* production of reuterin by *L. reuteri* has not been demonstrated. *Lactobacillus fermentum* is a closely related species based on a sequence analysis of the 16S ribosomal RNA gene (95% identity)[14] and on phenotypic properties, including being an obligate heterofermentative organism.[15] Previously, both these species were classified as a single species, *L. fermentum*, but were subsequently separated based primarily on DNA hybridization and GC content.[15] Despite many similar phenotypic characteristics, probiotic effects have primarily been observed for *L. reuteri*, with a few studies suggesting probiotic properties for *L. fermentum*.[16–18] To provide the framework for a comprehensive examination of the probiotic mechanisms of *L. reuteri*, we determined the complete genome sequences of the reuterin-producing *L. reuteri* JCM 1112^T, which was isolated from human feces, and *L. fermentum* IFO 3956, which was isolated from fermented plant material, and compared them to the genomes of other available *Lactobacillus* genomes.[19–26] Furthermore, we present *in vivo* evidence for the production of reuterin in the intestine of BALB/c mice.

2. Materials and methods

2.1. Genome sequencing

Lactobacillus reuteri JCM 1112^T, which was originally isolated from human feces, is a type of strain and was obtained from the Japan Collection of Microorganisms (JCM), Riken, Japan. *Lactobacillus fermentum* IFO 3956, which was originally isolated from fermented plant material, was obtained from the Institute for Fermentation, Osaka (IFO), Japan. Each strain was cultured in MRS broth (Oxoid) at 37°C for 18 h, and the genomic DNA was isolated and purified using standard genomic DNA affinity columns. The nucleotide sequences of the *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 genomes were determined by a whole-genome shotgun strategy. We constructed small-insert (2 kb) and large-insert (5 kb) genomic DNA libraries. Template DNA for sequencing was prepared by amplification of the DNA inserts using PCR with *Ex-Taq* (Takara Bio Inc.). Sequencing was subsequently carried out using the dye terminator strategy before precipitation with ethanol and running on ABI 3730xl (Applied Biosystems) and MegaBACE 4500 (GE Healthcare) capillary sequencers. We generated 27 456 reads (7.4-fold coverage) for JCM 1112^T and 35 136 reads (9.5-fold coverage) for IFO 3956, respectively, from both ends of the genomic clones. Data were assembled with the Phred-Phrap-Consed program,[27] and gaps were closed by direct sequencing of clones that spanned the gaps or of PCR products amplified with oligonucleotide primers designed to anneal to each end of neighboring contigs. The overall accuracy of the finished sequence was estimated to have an error rate of <1 per 10 000 bases (PHRAP score of ≥ 40).

2.2. Sequence analysis

An initial set of open reading frames (ORFs) that likely encode proteins was identified using Glimmer 2.0[28] and overlapped ORFs, and those shorter than 120 bp were eliminated. All predicted proteins were searched against a non-redundant protein database (nr, NCBI) using BLASTP. The start codon of each ORF was manually refined from BLASTP alignments. The tRNA genes were predicted by the tRNAscan-SE[29] and the rRNA genes were detected using BLASTN in a database of known bacterial RNAs. Functional classification of ORFs was conducted using the NCBI clusters of orthologous groups (COG) and LaCOG databases.[25,30] Protein domains were identified using the Pfam database.[31] The codon adaptation index for each protein was predicted using the cai program within the EMBOSS package[32] based on the codon usage preference of genes encoding ribosomal proteins. Sequences were aligned with ClustalW, and BLASTP was used for protein comparisons. Orthologous genes were identified as the reciprocal best matches using BLASTP. The data were parsed by custom Perl script. Genome sequences of *L. reuteri*

strains DSM 20016^T (accession no. CP000705) and 100-23 (accession no. AAPZ00000000), which were determined by Joint Genome Institute (JGI), were obtained from GenBank. The genome sequence data of *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 have been deposited in GenBank/DDBJ/EMBL under accession nos. AP007281 and AP008937, respectively.

2.3. Gnotobiotic mice assays

Lactobacillus reuteri JCM 1112^T and LR*ΔgupCDE* were cultured on MRS agar at 37°C for 3 days before pelleting and resuspending in phosphate-buffered saline (pH 7.2). BALB/c female mice (13 weeks old and 20–23 g) raised under germ-free conditions in the Gnotobiote Laboratory, University of Tokyo, were used.[33] Three germ-free mice were inoculated orally with 0.5 mL of a suspension of each bacterial strain (10^8 cfu/mL). The mice bacterial counts of JCM 1112^T and LR*ΔgupCDE* were 10^7 cfu/g feces. To detect *in vivo* reuterin production from *L. reuteri* JCM 1112^T in mono-associated mice, an intestinal loop assay was conducted under anesthetic.[34] The front and back of the cecum was bound with twine, and 100 μL of 2 M $^{13}C_3$-glycerol solution was injected. After 3 h, each mouse was killed by cervical vertebra dislocation and the cecal contents were collected.

2.4. Analysis of glycerol metabolism in L. reuteri JCM 1112^T using $^{13}C_3$-labeled glycerol and 2D-NMR

Lactobacillus reuteri JCM 1112^T and the mutant strain LR*ΔgupCDE* were grown in 30 mL serum vials containing 20 mL of MRS broth with 200 mM $^{13}C_3$-glycerol. Cultures were incubated in duplicate at 37°C for 10 h before 1 mL of each was sub-inoculated into new tubes for another 2.5 h incubation. An 0.1 mL aliquot of D_2O was then added for 2D-NMR analysis. To measure the *in vivo* glycerol metabolism, metabolites were extracted with 0.5 mL of CD_3OD from 1 g (wet weight) of JCM 1112^T and LR*ΔgupCDE* mono-associated murine cecal contents, as described previously.[35] All NMR spectra of CD_3OD-extracted sample solutions were recorded on a Bruker DRU-700 spectrometer equipped with a *Z*-axis gradient and a Bruker DRX-500 spectrometer equipped with a triple axis gradient. The temperature of the NMR samples was maintained at 298 K. A total of 128 complex f1 (^{13}C) and 1024 complex f2 (^1H) points were recorded with 32 scans per f1 increment for 2D-^{13}C-HSQC[36] and 2D-^{13}C-edited HSQC[37] spectra. The spectral widths were 12 000 and 8400 Hz for f1 and f2, respectively. In order to confirm signal assignments by ^3JHH connectivity's, 2D-^{13}C-HSQC-TOCSY spectra were recorded by a total of 64 complex f1 (^{13}C) and 1024 complex f2 (^1H) points using 32 scans per f1 increment.[38] To quantify the signal intensities, a Lorentzian-to-Gaussian window with a Lorentzian line width of 10 Hz and a Gaussian line width of 15 Hz was applied in both dimensions, prior to Fourier transformation. A fifth-order polynomial baseline correction was subsequently applied in the f1 dimension. The indirect dimension was zero-filled to 2048 points in the final data matrix. $^{13}C_3$-Glycerol was purchased from Cambridge Isotope Laboratories, Inc., and 1,3-propanediol and 3-hydroxypropionic acid were purchased from Tokyo Chemical Industry Co., Ltd.

3. Results and discussion

3.1 Genome features of L. reuteri JCM 1112^T and L. fermentum IFO 3956

The genomes of *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 each consist of single circular chromosomes of 2 039 414 and 2 098 685 bp, respectively, and both have no plasmids (Supplementary Fig. S1). The general features of both genome sequences are summarized in Table 1. The genomes of *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 contain 1820 and 1844 ORFs, respectively, and both shared 1265 ORFs. A high level of synteny is conserved between the two genomes (Fig. 1A). The close phylogenetic relationship is also evident from a comparative analysis of the ribosomal protein sequences from all lactobacilli, whose genomes have been

Table 1. General features of the *L. reuteri* and *L. fermentum* genomes

	L. reuteri JCM 1112^T	*L. fermentum* IFO 3956
Chromosome size (bp)	2 039 414	2 098 685
GC content (%)	38.9	51.5
Total ORFs	1820	1844
Functionally assigned	1211	1212
Conserved hypothetical	413	360
Non-conserved hypothetical	196	272
GC content at each codon position (%)		
First position	51.4	56.2
Second position	35.3	38.0
Third position	32.0	64.7
Coding density (%)	83.6	80.4
tRNAs	58	54
rRNA operons	6	5
Phage-related ORFs	53	24
Transposases	55	106
Group II introns	12	0

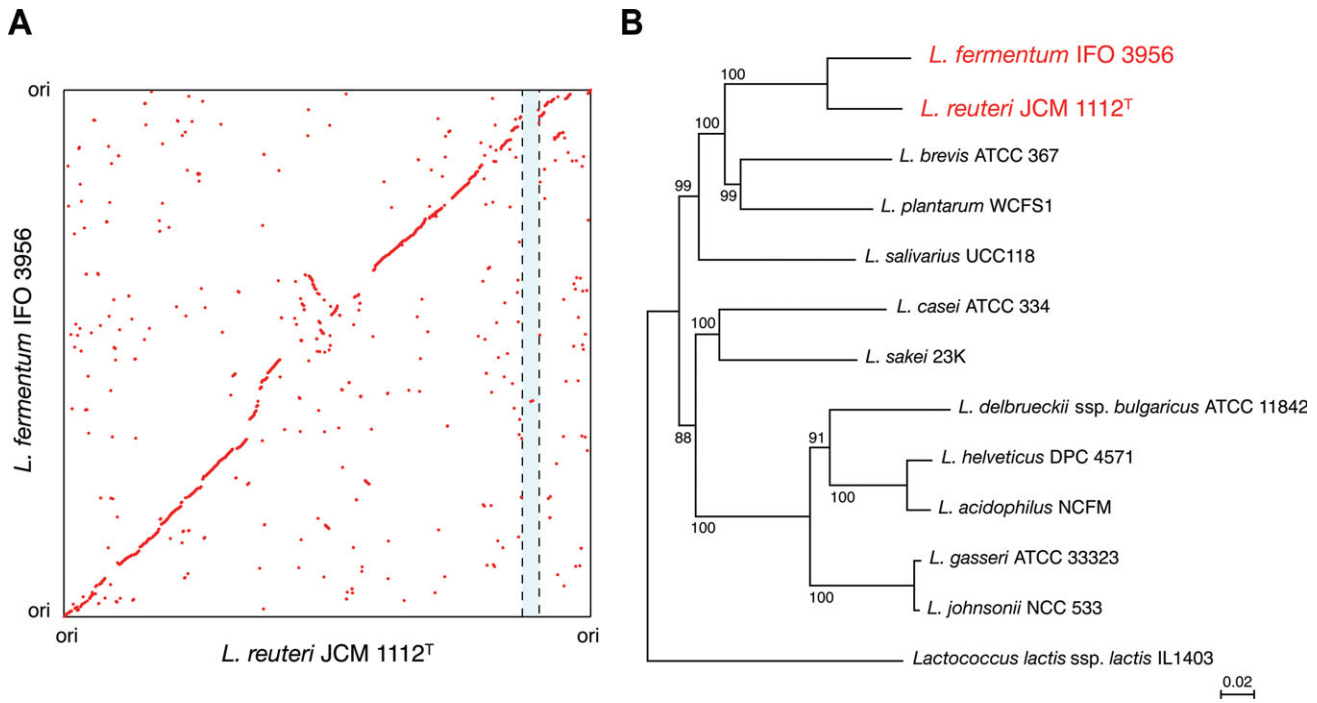

Figure 1. Genome-based phylogenetic analysis of *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956. (**A**) Synteny between the *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 chromosomes. Each dot represents an orthologous gene that was defined by bidirectional best hits based on BLASTP comparisons. Genome numbering was initiated at *dnaA* in both chromosomes. The shaded region indicates the location of the *pdu-cbi-cob-hem* cluster in the *L. reuteri* JCM 1112^T genome. (**B**) Phylogenetic relationships between the genomes of sequenced *Lactobacillus* species and other lactic acid bacteria, including *Lactococcus lactis*, inferred from 34 concatenated ribosomal protein amino acid sequences. The scale bar represents an evolutionary distance. Sequences were aligned with ClustalW with a bootstrap trial of 1000 and bootstrap values (%) are indicated at the nodes. An unrooted tree was generated using NJplot.

sequenced, and thus the two species constitute a distinct group (Fig. 1B). Despite the close phylogenic relationship of *L. reuteri* and *L. fermentum*, the *L. fermentum* IFO 3956 genome has a significantly higher GC content (52%) than that of the *L. reuteri* JCM 1112^T genome (39%) and the highest among the sequenced *Lactobacillus* genomes. This can be explained primarily by the GC content at the third codon position between *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 (32 versus 65%; Table 1).

During the preparation of this manuscript, the genome sequence (1 999 618 bp) of another *L. reuteri* strain DSM 20016^T was deposited in public databases, and this strain also produces reuterin. It was expected that there would be minimum sequence differences between the two strains, because both are derived from the same original isolate, *L. reuteri* F275.[39] However, an alignment of both genome sequences shows two large unique regions in *L. reuteri* JCM 1112^T that result in a 40-kb increase in its genome (Supplementary Table S1). This highlights the tendency of bacteria to undergo significant genome changes during regular laboratory culturing and transfers. One unique region (JCM 1112^T genome coordinates 442 995 to 451 429; unique region I) consists of 8435 bp and is flanked by a copy of IS*4* on either end (LAR_0380 and LAR_0385), and the remnant (155 bp) of IS*4* is present between orthologous genes (Lreu_0389 and Lreu_0390) at this position in the DSM 20016^T genome. It is difficult to imagine that a pure culture stock picks up additional DNA, and thus DSM 20016^T might have lost these genes. The unique region I contains the genes for enzymes involved in glycolysis (Supplementary Table S1). The other unique region (JCM 1112^T genome coordinates 1 064 161 to 1 094 397; unique region II) consists of 30 237 bp and is bordered by an IS*4* element (LAR_935) at one end and an IS*200* element (LAR_959) on the other, both of which are also present at this position (Lreu_1000 and Lreu_1004) in the DSM 20016^T genome. The unique region II contains the gene cluster encoding nitrate reductase subunits and the genes involved in molybdopterin biosynthesis (Supplementary Table S1). This gene cluster (LP_1473-1500) is also conserved in the genome of *Lactobacillus plantarum* WCFS1,[19] and it is one of the variable regions in *L. plantarum* strains.[40]

Biological roles were assigned to 1211 (67%) of the predicted ORFs for *L. reuteri* JCM 1112^T, and 1212 (66%) for *L. fermentum* IFO 3956. Obligately heterofermentative lactobacilli, such as *L. reuteri*, *L. fermentum*, and *Lactobacillus brevis*, produce CO2, ethanol, acetate, and lactate from metabolism of glucose,

whereas facultative heterofermentative and obligately homofermentative lactobacilli produce only lactate. The primary catabolic pathway for glucose metabolism in *Lactobacillus* is believed to be the glycolytic pathway. Surprisingly, the *pfk* gene encoding 6-phosphofructokinase is absent from the *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 genomes, and the *fba* gene for fructose-bisphosphate aldolase is also missing in IFO 3956. The *pfk* and *fba* genes are essential for the glycolytic pathway. The *L. brevis* genome also lacks the *pfk* and *fba* genes,[25] suggesting that *pfk* and *fba* are key genes distinguishing homo- and hetero-fermentation. On the other hand, both have the complete gene set for the pentose phosphate pathway. These findings indicate that *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 use the pentose phosphate pathway to metabolize glucose (Fig. 2). Of all the ORFs in both genomes, 181 (10%) in JCM 1112^T and 187 (10%) in IFO 3956 have no homolog in the other sequenced *Lactobacillus* genomes. Most of these genes encode conserved hypothetical proteins, but some with predicted functions are present. Genes encoding cobalamin biosynthesis are found only in JCM 1112^T (described later) and genes encoding allantoate amidohydrolase (LAF_0160), isocitrate dehydrogenase (LAF_0939), isopropylmalate isomerase (LAF_0963-0965), and nitric oxide reductase (LAF_1644) are found only in IFO 3956. A COG analysis showed that *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 contain fewer glycosidases and no sugar ATP-binding cassette transporter when compared with other lactobacilli (data not shown). JCM 1112^T contains the phosphoenolpyruvate:carbohydrate phosphotransferase systems (PTS) subunit IIC for cellobiose (LAR_0225) and galactitol (LAR_1559), but completely lacks other PTS enzyme II subunits. IFO 3956 has only two complete PTS enzyme II complexes for mannose (LAF_0391-0393) and sucrose (LAF_1579), but its gene-encoding mannose-6-phosphate isomerase is disrupted. *Lactobacillus reuteri* JCM 1112^T and *L. fermentum* IFO

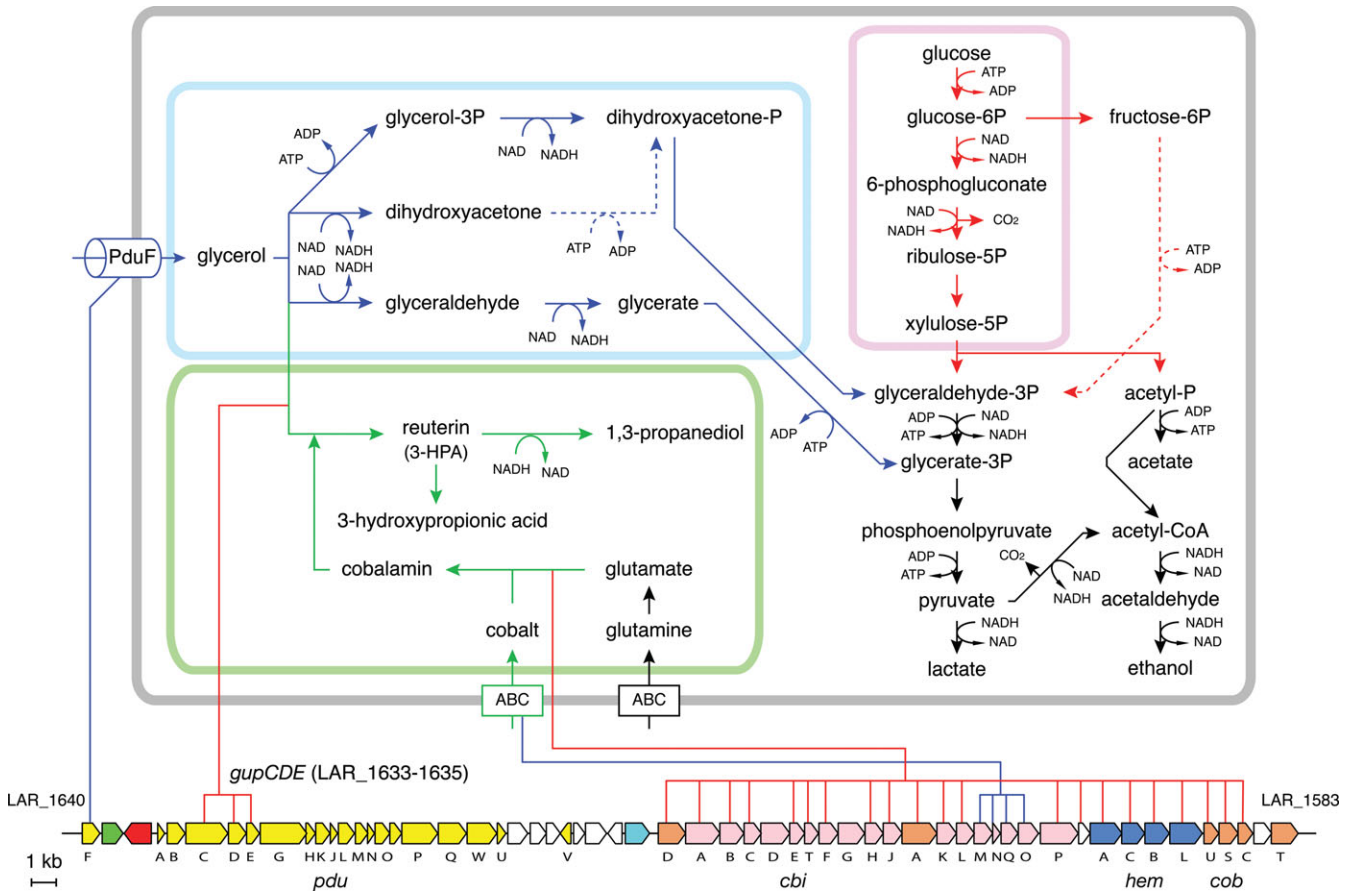

Figure 2. Proposed glycerol and glucose metabolic pathways in *L. reuteri* JCM 1112^T. The pathways are color coded as follows: blue, glycerol metabolism; green, biosynthesis of reuterin; red, glucose metabolism. Dashed lines indicate an unidentified enzyme in *L. reuteri* JCM 1112^T. The bottom part of the figure shows the structure of the *pdu-cbi-cob-hem* gene cluster in *L. reuteri* JCM 1112^T. Genes are depicted with arrows indicating the transcription direction with the following colors: yellow, *pdu* including *gupCDE* genes; pink, *cbi* genes; orange, *cob* genes; blue, *hem* genes; red, *pocR;* green, *eut*; sky blue, transposase gene; and white, other genes. Several lines connect corresponding genes between the pathway and the cluster (enzymes, red; transporters, blue).

3956 apparently have a reduced capacity to utilize carbohydrates compared with other lactobacilli.

3.2 Gene cluster for reuterin and cobalamin production in L. reuteri JCM 1112^T

Synthesis of reuterin is mediated by glycerol dehydratase (EC 4.2.1.30), which catalyzes the conversion of glycerol to 3-HPA, and this enzyme exists as a dimer of the three subunits α_2, β_2, and γ_2 that exhibit significant sequence and structural similarities to the corresponding subunits of propanediol dehydratase.[41] The glycerol dehydratase of some *L. reuteri* strains has been purified and characterized,[42] but not at the sequence level. JCM 1112^T contains three genes (LAR_1633-1635) with dehydratase subunit motifs (Pfam PF02286-02288) in the propanediol utilizing operon (LAR_1616-1640) (Fig. 2). The encoded proteins of these three genes showed sequence similarity to PduCDE (propanediol dehydratase) of *Salmonella typhimurium*[43] and thus the three genes were designated *gupCDE* (glycerol utilization gene candidates in the *pdu* operon) in JCM 1112^T. The *pdu* operon from *Lactobacillus collinoides* has been sequenced,[44] and it is also present in the *L. brevis* genome,[25] but the enzymatic activity has not been examined. Homologues for PduCDE have also been reported for *Lactobacillus hilgardii* and *Lactobacillus diolivorans*.[45] A comparison of the *pdu* operon in *L. reuteri* JCM 1112^T with *L. brevis* and *L. collinoides* reveals a conserved gene order (Supplementary Fig. S2). The encoded amino acid sequences of the *gupCDE* genes exhibit the highest identity to their homologues from *L. brevis* ATCC 367, showing 81, 66, and 57% identities, respectively (Supplementary Fig. S2). There is no synteny between the flanking region of the *pdu* operon of *L. reuteri* JCM 1112^T and *L. brevis* ATCC 367.

Both propanediol and glycerol dehydratase enzymes require cobalamin (vitamin B_{12}) as a cofactor.[41] JCM 1112^T possesses the gene sets (*cbi*, *cob*, and *hem*) for cobalamin biosynthesis, and these are located adjacent to the *pdu* operon (Fig. 2). The structure of this gene cluster represents a close association of two operons (*pdu* and *cbi-cob*), and this is probably a reflection of the cobalamin requirement for glycerol dehydratase activity.[41,46] These genes for cobalamin biosynthesis have not been identified in other lactobacilli including *L. brevis*. A comparison of this genetic locus with the corresponding locus in *L. fermentum* IFO 3956 and *L. plantarum* WCFS1 revealed that this *pdu-cbi-cob-hem* gene cluster comprising 58 ORFs (LAR_1583-1640), including 20 *pdu* genes (including *gupCDE*), 17 *cbi* genes, 6 *cob* genes, 4 *hem* genes, 1 *pocR* gene, 1 putative transposase gene, and 9 non-conserved ORFs, was apparently inserted into a locus that is common to all three bacteria (Fig. 3A). The *pocR* gene encodes a positive regulator of both the *pdu* and *cbi-cob* operons.[47] The average GC content (25%) at the third codon of the ORFs in this gene cluster is significantly lower than that (32%) of the remaining ORFs in the JCM 1112^T genome (Fig. 3A). In addition, there is a putative transposase gene (LAR_1612) at the boundary between the *pdu* and *cbi-cob-hem* operons (Fig. 2) and four IS elements within 20 kb of both the flanking regions of the *pdu-cbi-cob-hem* cluster in JCM 1112^T. These data strongly suggest that the *pdu-cbi-cob-hem* cluster in *L. reuteri* JCM 1112^T is a genomic island that has been acquired through lateral gene transfer. There is no direct repeat or tRNA gene in the flanking regions of this cluster, and the mechanism of the acquisition of this island is unclear. This finding is further supported from a comparison with the draft genome sequence data (total, 2.1 Mb) of *L. reuteri* 100-23, a rodent-specific strain, in which the *pdu-cbi-cob-hem* gene cluster is absent.

The *pdu-cbi-cob* gene cluster has been identified in the genomes of pathogenic bacteria such as *Salmonella* spp.,[48] *Listeria* spp.,[49] *Shigella sonnei*,[50] *Yersinia enterocolitica*,[51] and *Streptococcus sanguinis*[52] (Fig. 3B and Supplementary Fig. S2). However, the *hem* genes are clustered at different positions in the genomes of *Yersinia*, *Shigella*, and *Listeria*, and are scattered in the *Salmonella* genome. Only one gene (*eutJ*) of 12 ethanolamine utilizing genes (*eut*) in *Listeria* is encoded between *pocR* and *pduF* in JCM 1112^T. Given the different gene organization of the *pdu* and *hem* operons, as well as a different codon adaptation index (CAI)[53] between the *pdu* (0.55) and *cbi-cob/hem* (0.47) operons of JCM 1112^T, it suggests that the *pdu* and *cbi-cob/hem* operons may have been independently inserted in these genomes and these clusters in *L. reuteri* JCM 1112^T may have been generated by at least two independent insertion events rather than by a single event. However, the possibility that the *pdu-cbi-cob* gene cluster might have been formed prior to the integration into the JCM 1112^T genome cannot be excluded.

3.3 Production of reuterin and cobalamin by L. reuteri JCM 1112^T

The *gupCDE* genes from *L. reuteri* JCM 1112^T were cloned in *Escherichia coli* to experimentally examine their encoded glycerol dehydratase activity. Some glycerol dehydratase activity was detected in *E. coli* carrying *gupCDE*, but not in *E. coli* derivatives containing the individual genes (Supplementary Table S2), indicating that the gene products of *gupCDE* are the subunits of glycerol dehydratase. Wild-type *L. reuteri* JCM 1112^T and a *gupCDE* knockout mutant (LRΔ*gupCDE*) (Supplementary Fig. S3), grown in MRS broth

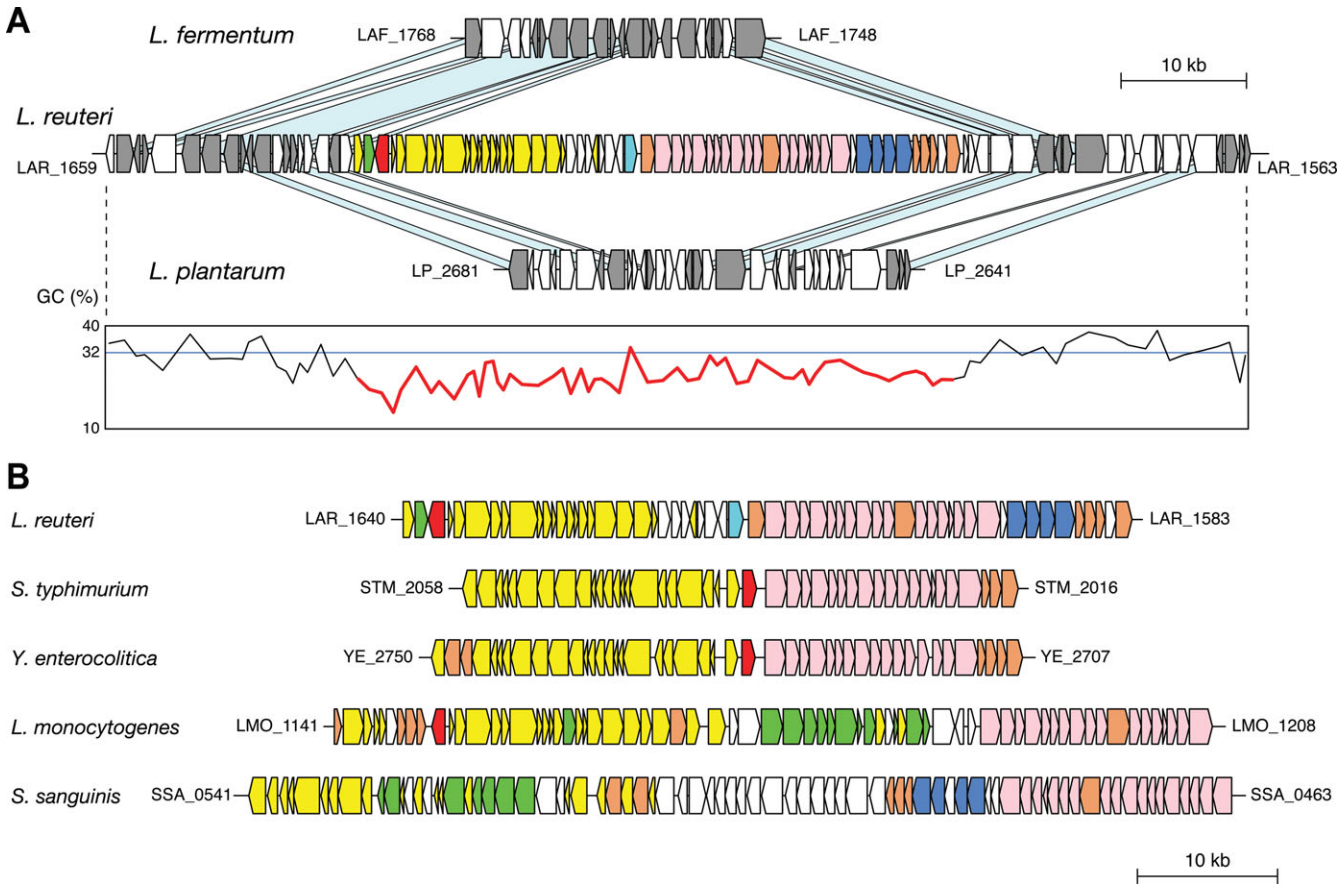

Figure 3. (**A**) A comparison of the genomic location that contains the *pdu-cbi-cob-hem* gene cluster of *L. reuteri* JCM 1112^T (center) with the corresponding location of *L. fermentum* IFO 3956 (upper) and *L. plantarum* WCFS1 (lower). Genes in the *pdu-cbi-cob-hem* gene cluster are depicted by arrows indicating the transcription direction with the same color codes as described in the Fig. 2. Genes conserved between the three genomes are colored gray and light blue bars indicate orthologous regions. The GC content at the third codon position of the ORFs in *L. reuteri* JCM 1112^T is indicated under each ORF. Red lines represent the GC content at the third codon position of the ORFs in the *pdu-cbi-cob-hem* cluster (LAR_1583-1640). The blue horizontal line indicates the average GC content (32%) at the third codon position of the remaining ORFs in the *L. reuteri* JCM 1112^T genome excluding the *pdu-cbi-cob-hem* gene cluster. (**B**) The *pdu-cbi-cob* gene cluster arrangement in *L. reuteri* JCM 1112^T, *S. typhimurium* LT2, *L. monocytogenes* EGD-e, *Y. enterocolitica* subsp. *enterocolitica* 8081, and *S. sanguinis* SK36 are shown using the same color coding as described in (A).

supplemented with glucose and $^{13}C_3$-glycerol, were analyzed using 2D-NMR to determine whether the proteins encoded by *gupCDE* can convert glycerol to 3-HPA. After 7.5 h incubation, 3-HPA was detected in JCM 1112^T, but not in LR*ΔgupCDE* under the same conditions (Supplementary Fig. S4). These results demonstrate that the glycerol dehydratase encoded by *gupCDE* in JCM 1112^T can convert glycerol to reuterin, consistent with previous studies using *L. reuteri* 1063.[9]

Lactobacillus reuteri JCM 1112^T also possesses the gene (LAR_0029) encoding 1,3-propanediol dehydrogenase (Fig. 2). While 1,3-propanediol was detected in JCM 1112^T after 2.5 h incubation (Supplementary Fig. S4), 3-HPA was first detected after 7.5 h incubation during culture in the MRS broth. These results reveal that 1,3-propanediol accumulated faster and more abundantly through the fermentation of glycerol than 3-HPA in *L. reuteri* cells. Negative effects of accumulated 1,3-propanediol together with a lower pH and an increase in NAD^+ may be involved in the subsequent accumulation and extracellular release of 3-HPA by *L. reuteri*, as was previously observed in the fermentation of glycerol involving 1,3-propanediol dehydrogenase in *Enterobacter agglomerans*.[54] JCM 1112^T produced neither ^{13}C-labeled lactic acid nor ^{13}C-labeled ethanol, which are fermentation end products (Supplementary Fig. S4). JCM 1112^T cannot be grown in MRS broth in which glycerol is the sole carbon source (data not shown). Taken together, these observations indicate that fermentation of glycerol for energy and growth did not occur in *L. reuteri*, which is consistent with the lack of the gene for dihydroxyacetone kinase in JCM 1112^T for conversion of glycerol to dihydroxyacetone phosphate prior to metabolism via the Embden–Meyerhof–Parnas pathway (Fig. 2).

As discussed above, JCM 1112^T contains the *cbi-cob-hem* gene cluster for cobalamin biosynthesis. Cobalamin production by *L. reuteri* was analyzed as described previously[55] and was found to be present in the cell extract of JCM 1112^T (Supplementary Table S3). This confirmed that JCM 1112^T has the ability to produce both reuterin and cobalamin, consistent with the predicted gene products of the *pdu-cbi-cob-hem* gene cluster. It was previously demonstrated that adenosylcobalamin has a higher affinity for the glycerol dehydratase of *L. reuteri* than that of other *Lactobacillus* species.[56] A high affinity of glycerol dehydratase for cobalamin may accelerate reuterin production by *L. reuteri*.

3.4. *In vivo production of reuterin by L. reuteri JCM 1112^T*

In vivo reuterin production by *L. reuteri* JCM 1112^T was investigated using gnotobiotic mice mono-associated with JCM 1112^T or its mutant derivative LR*ΔgupCDE*. The intestinal loop assay of mono-associated mice was conducted 7 days after injection of a $^{13}C_3$-glycerol solution into the murine cecum. The results revealed that reuterin was detected in samples taken from the cecum of gnotobiotic mice mono-associated with JCM 1112^T, but not in mice mono-associated with LR*ΔgupCDE* (Fig. 4). More rapid and abundant accumulation of 1,3-propanediol than that of 3-HPA was also observed *in vivo*, similar to the *in vitro* analysis (Supplementary Fig. S4). These results clearly demonstrated that glycerol dehydratase in *L. reuteri* converts glycerol into reuterin *in vivo*, supporting the hypothesis that the proposed probiotic effects of reuterin may occur in the intestine when glycerol is present. An analysis of the feces of untreated mice using NMR spectrometry and gas chromatography revealed 7–10 mM glycerol was present, suggesting the presence of glycerol in the intestine (data not shown). This is the first demonstration that the antimicrobial substance (reuterin) produced by *L. reuteri* can be produced *in vivo* in the gut under physiological conditions. This supports the genomic analysis as the genomic island encoding this phenotype, in this intestinal microbe, would likely not have been acquired unless it conferred a competitive advantage.

3.5. *Conclusions*

The complete genome sequences of *L. reuteri* JCM 1112^T and *L. fermentum* IFO 3956 were determined. A comparative genome analysis suggested that *L. reuteri* JCM 1112^T contains a specific gene cluster for reuterin and cobalamin production, and that the gene cluster is encoded on a genomic island that was acquired through lateral gene transfer. The *pdu-cbi-cob* gene cluster has also been identified in the genomes of some pathogenic bacteria, and previous studies have suggested that *S. typhimurium* acquired the *pdu-cbi-cob* gene cluster by horizontal transfer to enable it to degrade propanediol.[43] As the *pdu*, *cbi*, and *cob* operons are not universally present in all salmonellae,[57] it supports the hypothesis that the genes for propanediol utilization (*pdu*) and cobalamin biosynthesis (*cbi-cob*) may have been acquired via horizontal gene in response to environmental evolutionary adaptation. Genes for anaerobic

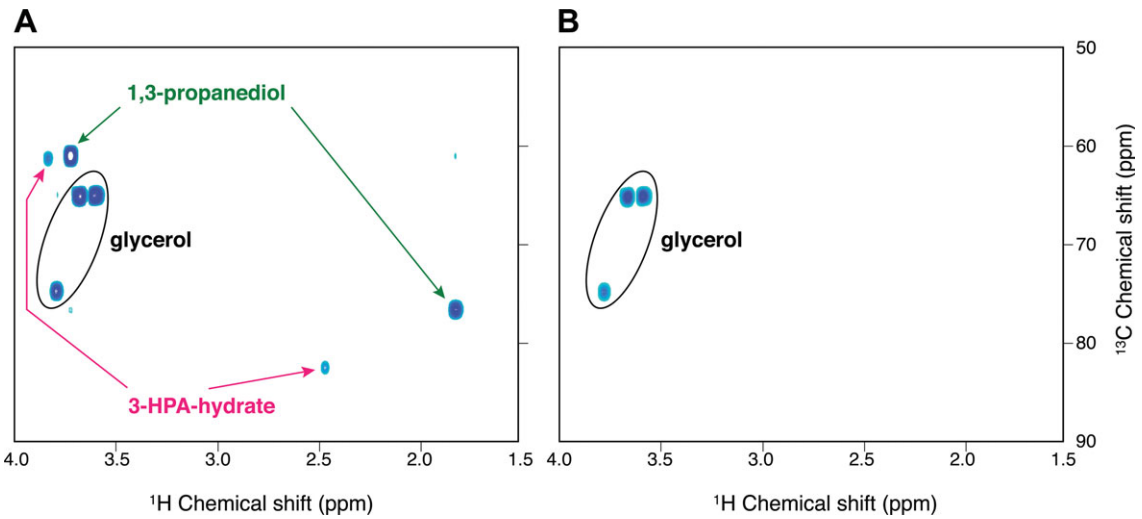

Figure 4. *In vivo* detection of 3-HPA-hydrate derived from $^{13}C_3$-glycerol in the cecal contents of mice colonized by wild-type *L. reuteri* JCM 1112^T (**A**) and its mutant derivative, LR*ΔgupCDE* (**B**), as detected by 2D ^{1}H, ^{13}C-HSQC NMR. δ^1H (ppm) and δ^{13}C (ppm) of 3-HPA-hydrate 2 (Hh-2) and 3-HPA-hydrate 3 (Hh-3) were observed at 2.43 and 42.6, and 3.79 and 61.6 at pH 7.0, respectively. Although the spot derived from 3-HPA-hydrate 1 (Hh-1) in ^{1}H chemical shift was detected, the spot was not shown in (A), because the range of ^{1}H chemical shift showed between 4.0 to 1.5 ppm. See Supplementary Figure S5 for details.

metabolism of glycerol are located in the dihydroxyacetone (*dha*) regulon, whereas genes for the anaerobic degradation of propanediol are part of the *pdu* operon.[41] The dehydratases of *L. collinoides* and *L. brevis* are general diol dehydratases (EC 4.2.1.28) rather than glycerol-specific dehydratases,[56] but the dehydratase of *L. reuteri* has a higher affinity for glycerol than for propanediol,[42] thus making it more efficient for conversion of glycerol to reuterin. The gene products of the homologues of *pduCDE* might determine properties of the *pdu-cbi-cob* gene cluster, and this cluster including *gupCDE* might provide a selective advantage to *L. reuteri*. Given the positive attributes of this phenotype, the *pdu-cbi-cob-hem* gene cluster might contribute to the probiotic properties of lactic acid bacteria. Genomic islands involved in pathogenesis or symbiosis are believed to be significant in the evolution of different bacterial species. In this study, the *L. reuteri* genome has a genomic island that may be important in the evolution of *L. reuteri* strains as health promoting bacteria in the human gut.

Acknowledgments: We are grateful to the consortium members of Azabu University Academic Frontier Project. We thank A. Yamashita, K. Furuya, C. Yoshino, H. Inaba, Y. Yamashita, A. Tamura, and N. Ito (Kitasato University), E. Okada and Y. Arita (Hitachi High-Tech Science Systems Co., Ltd.), and J.H. Lee (University of Minnesota) for technical assistance. We also thank DOE JGI for use of unpublished data of *L. reuteri* 100-23 and *L. reuteri* DSM 20016^T.

Supplementary Data: Supplementary data are available online at www.dnaresearch.oxfordjournals.org.

Funding

This research was supported by the 'Academic Frontier' Project for Private Universities: Matching Fund Subsidy, 2002-2006 (Azabu University), the Grants-in-Aid of Scientific Research (C) (18580275), Grant-in-Aid for Scientific Research on Priority Areas "Comprehensive Genomics" (17020007), and Grant-in-Aid for Encouragement of Young Scientists (B) (17710191) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. Funding for open access charge: the 'Academic Frontier' Project for Private Universities: Matching Fund Subsidy (Azabu University).

References

1. Casas, I. A. and Dobrogosz, W. J. 1997, *Lactobacillus reuteri*: an overview of a new probiotic for humans and animals, *Microecol. Therap.*, **26**, 221–231.
2. Casas, I. A. and Dobrogosz, W. J. 2000, Validation of the probiotic concept: *Lactobacillus reuteri* confers broad-spectrum protection against disease in humans and animals, *Microb. Ecol. Health Dis.*, **12**, 247–285.
3. Shornikova, A. V., Casas, I. A., Isolauri, E., Mykkänen, H. and Vesikari, T. 1997, *Lactobacillus reuteri* as a therapeutic agent in acute diarrhea in young children, *J. Pediatr. Gastroenterol. Nutr.*, **24**, 399–404.
4. Valeur, N., Engel, P., Carbajal, N., Connolly, E. and Ladefoged, K. 2004, Colonization and immunomodulation by *Lactobacillus reuteri* ATCC 55730 in the human gastrointestinal tract, *Appl. Environ. Microbiol.*, **70**, 1176–1181.
5. Chang, Y. H., Kim, J. K., Kim, H. J., Kim, W. Y., Kim, Y. B. and Park, Y. H. 2001, Selection of a potential probiotic *Lactobacillus* strain and subsequent *in vivo* studies, *Antonie van Leeuwenhoek*, **80**, 193–199.
6. Rosenfeldt, V., Michaelsen, K. F., Jakobsen, M., et al. 2002, Effect of probiotic *Lactobacillus* strains on acute diarrhea in a cohort of nonhospitalized children attending day-care centers, *Pediatr. Infect. Dis. J.*, **21**, 417–419.
7. Weizman, Z., Asli, G. and Alsheikh, A. 2005, Effect of a probiotic infant formula on infections in child care centers: comparison of two probiotic agents, *Pediatrics*, **115**, 5–9.
8. Niv, E., Naftali, T., Hallak, R. and Vaisman, N. 2005, The efficacy of *Lactobacillus reuteri* ATCC 55730 in the treatment of patients with irritable bowel syndrome: a double blind, placebo-controlled, randomized study, *Clin. Nutr.*, **24**, 925–931.
9. Talarico, T. L., Casas, I. A., Chung, T. C. and Dobrogosz, W. J. 1988, Production and isolation of reuterin: a growth inhibitor produced by *Lactobacillus reuteri*, *Antimicrob. Agents Chemother.*, **32**, 1854–1858.
10. Chung, T. C., Axelsson, L., Lindgren, S. E. and Dobrogosz, W. J. 1989, *In vitro* studies on reuterin synthesis by *Lactobacillus reuteri*, *Microb. Ecol. Health Dis.*, **2**, 137–144.
11. Talarico, T. L. and Dobrogosz, W. J. 1989, Chemical characterization of an antimicrobial substance produced by *Lactobacillus reuteri*, *Antimicrob. Agents Chemother.*, **33**, 674–679.
12. Lüthi-Peng, Q., Schärer, S. and Puhan, Z. 2002, Production and stability of 3-hydroxypropionaldehyde in *Lactobacillus reuteri*, *Appl. Microbiol. Biotechnol.*, **60**, 73–80.
13. Vollenweider, S. and Lacroix, C. 2004, 3-Hydroxypropionaldehyde: applications and perspectives of biotechnological production, *Appl. Microbiol. Biotechnol.*, **64**, 16–27.
14. Ennahar, S., Cai, Y. and Fujita, Y. 2003, Phylogenetic diversity of lactic acid bacteria associated with paddy rice silage as determined by 16S ribosomal DNA analysis, *Appl. Environ. Microbiol.*, **69**, 444–451.
15. Klein, G., Pack, A., Bonaparte, C. and Reuter, G. 1998, Taxonomy and physiology of probiotic lactic acid bacteria, *Int. J. Food Microbiol.*, **41**, 103–125.
16. Strompfová, V., Marcináková, M., Simonová, M., Bogovic-Matijasić, B. and Lauková, A. 2006, Application of potential probiotic *Lactobacillus fermentum* AD1 strain in healthy dogs, *Anaerobe*, **12**, 75–79.

17. Peran, L., Camuesco, D., Comalada, M., et al. 2006, *Lactobacillus fermentum*, a probiotic capable to release glutathione, prevents colonic inflammation in the TNBS model of rat colitis, *Int. J. Colorectal Dis.*, **21**, 737–746.
18. Geier, M. S., Butler, R. N., Giffard, P. M. and Howarth, G. S. 2007, *Lactobacillus fermentum* BR11, a potential new probiotic, alleviates symptoms of colitis induced bydextran sulfate sodium (DSS) in rats, *Int. J. Food Microbiol.*, **114**, 267–274.
19. Kleerebezem, M., Boekhorst, J., van Kranenburg, R., et al. 2003, Complete genome sequence of *Lactobacillus plantarum* WCFS1, *Proc. Natl. Acad. Sci. USA*, **100**, 1990–1995.
20. Pridmore, R. D., Berger, B., Desiere, F., et al. 2004, The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533, *Proc. Natl. Acad. Sci. USA*, **101**, 2512–2517.
21. Altermann, E., Russell, W. M., Azcarate-Peril, M. A., et al. 2005, Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM, *Proc. Natl. Acad. Sci. USA*, **102**, 3906–3912.
22. Chaillou, S., Champomier-Vergès, M. C., Cornet, M., et al. 2005, Complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K, *Nat. Biotechnol.*, **23**, 1527–1533.
23. Claesson, M. J., Li, Y., Leahy, S., et al. 2006, Multireplicon genome architecture of *Lactobacillus salivarius*, *Proc. Natl. Acad. Sci. USA*, **103**, 6718–6723.
24. van de Guchte, M., Penaud, S., Grimaldi, C., et al. 2006, The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution, *Proc. Natl. Acad. Sci. USA*, **103**, 9274–9279.
25. Makarova, K., Slesarev, A., Wolf, Y., et al. 2006, Comparative genomics of the lactic acid bacteria, *Proc. Natl. Acad. Sci. USA*, **103**, 15611–15616.
26. Callanan, M., Kaleta, P., O'Callaghan, J., et al. 2008, Genome sequence of *Lactobacillus helveticus*, an organism distinguished by selective gene loss and insertion sequence element expansion, *J. Bacteriol.*, **190**, 727–735.
27. Gordon, D., Desmarais, C. and Green, P. 2001, Automated finishing with autofinish, *Genome Res.*, **11**, 614–625.
28. Delcher, A. L., Harmon, D., Kasif, S., White, O. and Salzberg, S. L.. 1999, Improved microbial gene identification with GLIMMER, *Nucleic Acids Res.*, **27**, 4636–4641.
29. Lowe, T. and Eddy, S. R. 1997, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, *Nucleic Acids Res.*, **25**, 955–964.
30. Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., et al. 2001, The COG database: new developments in phylogenetic classification of proteins from complete genomes, *Nucleic Acids Res.*, **29**, 22–28.
31. Bateman, A., Coin, L., Durbin, R., et al. 2004, The Pfam protein families database, *Nucleic Acids Res.*, **32**, 138–141.
32. Rice, P., Longden, I. and Bleasby, A. 2000, EMBOSS: the European molecular biology open software suite, *Trends Genet.*, **16**, 276–277.
33. Tamura, M., Hirayama, K. and Itoh, K. 1996, Comparison of colonic bacterial enzymes in gnotobiotic mice monoassociated with different intestinal bacteria, *Microb. Ecol. Health Dis.*, **9**, 287–294.
34. Sakaguchi, T., Sakaguchi, S., Nakamura, I., Kudo, Y., Ichiman, Y. and Yoshida, K. 1988, Positive reaction in mouse ligated intestinal loop assay with nonenterotoxigenic and nonhemolytic strains of *Staphylococcus aureus*, *J. Clin. Microbiol.*, **26**, 600–601.
35. Kikuchi, J. and Hirayama, T. 2006, Practical aspects of stable isotope labeling of higher plants for a heteronuclear multi-dimensional NMR-based metabolomics, *Methods Mol. Biol.*, **358**, 273–286.
36. Vuister, G. W. and Bax, A.. 1992, Resolution enhancement and spectral editing of uniformly C-13-enriched proteins by homonuclear broad-band C-13 decoupling, *J. Magn. Reson.*, **98**, 428–435.
37. Parrella, F., Sanchez-Ferrando, A. and Virgili, A. 1997, Improved sensitivity in gradient-based 1D and 2D multiplicity-edited HSQC experiments, *J. Magn. Reson.*, **126**, 274–277.
38. Schmieder, P., Kurz, M. and Kessler, H. 1991, Determination of heteronuclear long-range couplings to heteronuclei in natural abundance by two- and three-dimensional NMR spectroscopy, *J. Biomol. NMR*, **1**, 403–420.
39. Kandler, O., Stetter, K. O. and Kohl, R. 1980, *Lactobacillus reuteri* sp. nov., a new species of heterofermentative lactobacilli, Zentbl. Bakteriol. Mikrobiol. Hyg. I Abt. Orig., ***C1***, 264–269.
40. Molenaar, D., Bringel, F., Schuren, F. H., et al. 2005, Exploring *Lactobacillus plantarum* genome diversity by using microarrays, *J. Bacteriol.*, **187**, 6119–6127.
41. Daniel, R., Bobik, T. A. and Gottschalk, G. 1998, Biochemistry of coenzyme B_{12}-dependent glycerol and diol dehydratases and organization of the encoding genes, *FEMS Microbiol. Rev.*, **22**, 553–566.
42. Talarico, T. L. and Dobrogosz, W. J. 1990, Purification and characterization of glycerol dehydratase from *Lactobacillus reuteri*, *Appl. Environ. Microbiol.*, **56**, 1195–1197.
43. Bobik, T. A., Xu, Y., Jeter, R. M., et al. 1997, Propanediol utilization genes (*pdu*) of *Salmonella typhimurium*: three genes for the propanediol dehydratase, *J. Bacteriol.*, **179**, 6633–6639.
44. Sauvageot, N., Muller, C., Hartke, A., Auffray, Y. and Laplace, J. M. 2002, Characterisation of the diol dehydratase pdu operon of *Lactobacillus collinoides*, *FEMS Microbiol. Lett.*, **209**, 69–74.
45. Gorga, A., Claisse, O. and Lonvaud-Funel, A. 2002, Organisation of the genes encoding glycerol dehydratase of *Lactobacillus collinoides, Lactobacillus hilgardii* and *Lactobacillus diolivorans*, *Sci. Aliments*, **22**, 151–160.
46. Lawrence, J. G. 2003, Gene organization: selection, selfishness, and serendipity, *Annu. Rev. Microbiol.*, **57**, 419–440.
47. Bobik, T. A., Ailion, M. and Roth, J. R. 1992, A single regulatory gene integrates control of vitamin B_{12} synthesis and propanediol degradation, *J. Bacteriol.*, **174**, 2253–2266.
48. McClelland, M., Sanderson, K. E., Spieth, J., et al. 2001, Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2, *Nature*, **413**, 852–856.

49. Glaser, P., Frangeul, L., Buchrieser, C., et al. 2001, Comparative genomics of *Listeria* species, *Science*, **294**, 849–852.
50. Yang, F., Yang, J., Zhang, X., et al. 2005, Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery, *Nucleic Acids Res.*, **33**, 6445–6458.
51. Thomson, N. R., Howard, S., Wren, B. W., et al. 2006, The complete genome sequence and comparative genome analysis of the high pathogenicity *Yersinia enterocolitica* strain 8081, *PLoS Genet.*, **2**, pe206.
52. Xu, P., Alves, J. M., Kitten, T., et al. 2007, Genome of the opportunistic pathogen *Streptococcus sanguinis*, *J. Bacteriol.*, **189**, 3166–3175.
53. Sharp, P. M. and Li, W. H. 1987, The codon adaptation index: a measure of directional synonymous codon usage bias, and its potential applications, *Nucleic Acids Res.*, **15**, 1281–1295.
54. Barbirato, F., Larguier, A., Conte, T., Astruc, S. and Bories, A. 1997, Sensitivity to pH, product inhibition and inhibition by NAD^+ of 1,3-propanediol dehydrogenase purified from *Enterobacter agglomerans* CNCM 1210, *Arch. Microbiol.*, **168**, 160–163.
55. Taranto, M. P., Vera, J. L., Hugenholtz, J., de Valdez, G. F. and Sesma, F. 2003, *Lactobacillus reuteri* CRL1098 produces cobalamin, *J. Bacteriol.*, **185**, 5643–5647.
56. Sauvageot, N., Pichereau, V., Louarme, L., Hartke, A., Auffray, Y. and Laplace, J. M. 2002, Purification, characterization and subunits identification of the diol dehydratase of *Lactobacillus collinoides*, *Eur. J. Biochem.*, **269**, 5731–5737.
57. Porwollik, S., Wong, R.M. and McClelland, M. 2002, Evolutionary genomics of *Salmonella*: gene acquisitions revealed by microarray analysis, *Proc. Natl. Acad. Sci. USA*, **99**, 8956–8961.