Applied and Environmental Microbiology, Aug. 2008, p. 4610–4625 Vol. 74, No. 15
0099-2240/08/$08.00+0 doi:10.1128/AEM.00054-08

Analysis of the Genome Sequence of *Lactobacillus gasseri* ATCC 33323 Reveals the Molecular Basis of an Autochthonous Intestinal Organism▽†

M. Andrea Azcarate-Peril,[1] Eric Altermann,[1]‡ Yong Jun Goh,[1] Richard Tallon,[1] Rosemary B. Sanozky-Dawes,[1] Erika A. Pfeiler,[1,2] Sarah O'Flaherty,[1] B. Logan Buck,[1]§ Alleson Dobson,[1]‖ Tri Duong,[1]# Michael J. Miller,[1]†† Rodolphe Barrangou,[1]‡‡ and Todd R. Klaenhammer[1]*

Department of Food, Bioprocessing, and Nutrition Sciences and Southeast Dairy Foods Research Center, North Carolina State University, Raleigh, North Carolina 27695,[1] and Genomics Sciences Graduate Program, North Carolina State University, Raleigh, North Carolina 27695[2]

Received 7 January 2008/Accepted 29 February 2008

This study presents the complete genome sequence of *Lactobacillus gasseri* ATCC 33323, a neotype strain of human origin and a native species found commonly in the gastrointestinal tracts of neonates and adults. The plasmid-free genome was 1,894,360 bp in size and predicted to encode 1,810 genes. The GC content was 35.3%, similar to the GC content of its closest relatives, *L. johnsonii* NCC 533 (34%) and *L. acidophilus* NCFM (34%). Two identical copies of the prophage LgaI (40,086 bp), of the Sfi11-like *Siphoviridae* phage family, were integrated tandomly in the chromosome. A number of unique features were identified in the genome of *L. gasseri* that were likely acquired by horizontal gene transfer and may contribute to the survival of this bacterium in its ecological niche. *L. gasseri* encodes two restriction and modification systems, which may limit bacteriophage infection. *L. gasseri* also encodes an operon for production of heteropolysaccharides of high complexity. A unique alternative sigma factor was present similar to that of *B. caccae* ATCC 43185, a bacterial species isolated from human feces. In addition, *L. gasseri* encoded the highest number of putative mucus-binding proteins (14) among lactobacilli sequenced to date. Selected phenotypic characteristics that were compared between ATCC 33323 and other human *L. gasseri* strains included carbohydrate fermentation patterns, growth and survival in bile, oxalate degradation, and adhesion to intestinal epithelial cells, in vitro. The results from this study indicated high intraspecies variability from a genome encoding traits important for survival and retention in the gastrointestinal tract.

Lactobacillus gasseri ATCC 33323/DSM 20243, previously known as "F. Gasser 63 AM," is a strain of human origin and a normal inhabitant of the mouths, intestines, feces, and vaginas of juveniles and adults (64). The bacterium is an obligate saccharoclastic, homofermentative organism, with an optimum growth at 35 to 38°C, and forms small rods with rounded ends from 0.6 to 0.8 by 3 to 5 μm in size. *L. gasseri* is considered one of the true autochthonous species of the human intestinal "probiome," defined here as commensal intestinal bacteria considered to have a beneficial influence on human health. Depending on consumption habits and geographic location, *L. gasseri* has been determined to be one of the *Lactobacillus* species native to the human gastrointestinal tract (GIT) of neonates (102) and adults (85). *L. gasseri* has also been described as a common member of the oral *Lactobacillus* biota (72, 91), and it has been proposed that the oral cavity acts as a reservoir and source of intestinal lactobacilli (28). Moreover, it has been proposed that the native *Lactobacillus* and *Bifidobacterium* microflora remains stable for life in a human being (85), although other transient species can temporarily alter the human microbiome "fingerprint" when fed in high numbers.

The structure and composition of the GIT microbiome reflects natural selection at both microbial and host levels, in a complex and delicate symbiotic state. In fact, it has been proposed that the microbiota acts as a multifunctional organ that contributes to essential human functions, such as immunomodulation and digestion (104). The GIT is sterile at birth but colonization begins immediately and is influenced by the infant diet, hygiene level, and other factors (38). Colonization of the stomach and proximal small intestine is limited due to the presence of acid, bile, and pancreatic secretions, with bacterial numbers ranging from 10^1 to 10^3 CFU/ml. Bacterial density

* Corresponding author. Mailing address: Department of Food, Bioprocessing, and Nutrition Sciences, North Carolina State University, Box 7624, Raleigh, NC 27695. Phone: (919) 515-2972. Fax: (919) 513-0014. E-mail: klaenhammer@ncsu.edu.

‡ Present address: AgResearch Limited, Grasslands Research Centre, Tennent Drive, Private Bag 11008, Palmerston North, New Zealand.

§ Present address: 575 Little Creek Rd., Banner Elk, NC 28604.

‖ Present address: Teagasc, Moorepark Food Research Centre, Fermoy, Co. Cork, and Department of Microbiology and Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland.

Present address: School of Molecular Biosciences, Washington State University, Pullman, WA 99164.

†† Present address: Department of Food Science and Human Nutrition, University of Illinois Urbana-Champaign, Urbana, IL 61801.

‡‡ Present address: Danisco USA, Inc., 3329 Agriculture Drive, Madison, WI 53716.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 6 June 2008.

increases in the distal small intestine (10^4 to 10^7 CFU/ml) to reach its maximum in the colon (10^{11} to 10^{12} CFU/ml) (77). *L. gasseri* has been regarded as a common autochthonous lactobacilli in the jejunum, as well as the ileum (85).

Before 1980, *L. gasseri* was routinely classified as "*L. acidophilus*" since morphologically it differs only slightly from *L. acidophilus* and cannot be distinguished from *L. acidophilus* by the classical taxonomic characteristics, such as carbohydrate utilization, lactic acid isomer produced, etc. (64). Originally, the *L. acidophilus* group was isolated from infant feces by Moro in 1900 and named "*Bacillus acidophilus*." Later, "*Bacillus acidophilus*" was included in the genus *Lactobacillus*. In 1980, *L. gasseri* was differentiated by DNA/DNA hybridization patterns from *L. acidophilus* and named after Francis Gasser, who studied lactate dehydrogenases of *Lactobacillus* species (44).

Today, there are 479 draft-phase or completed genomes deposited in the phylum *Firmicutes* in the Bacterial Genome Database at the National Center for Biotechnology Information (NCBI). A total of 173 sequences belong to the order *Lactobacillales*; 45 are of the family *Lactobacillaceae*, with 43 of them of the genus *Lactobacillus*. The genome of *L. gasseri* ATCC 33323 was sequenced by the Department of Energy-Joint Genome Institute in collaboration with the Lactic Acid Bacteria Genomics Consortium (LABGC) (56). This endeavor included sequencing of the genomes of 11 lactic acid bacteria (LAB) that are now publicly available at http://genome.jgi-psf.org/mic_home.html. In addition, the genome sequences of *L. plantarum* (61), *L. johnsonii* (83), *L. acidophilus* (4), *L. sakei* (23), *L. salivarius* (26), *L. delbrueckii* subsp. *bulgaricus* (97), and *L. helveticus* (22) have been published in the last 4 years. The LABGC project culminated with the analysis of Makarova et al. (68) that compared the genome sequences of *L. gasseri*, *Lactobacillus brevis*, *Pediococcus pentosaceus*, *Lactococcus lactis* subsp. *cremoris*, *Streptococcus thermophilus*, *Oenococcus oeni*, *Leuconostoc mesenteroides*, *L. casei*, and *L. delbrueckii* subsp. *bulgaricus*. In silico analyses of similarities and differences at the species level within the *Lactobacillus* group revealed extensive similarities at DNA and protein levels between *L. johnsonii* NCC533 and *L. gasseri* ATCC 33323 (16).

We sought here to present the annotated genome sequence of *L. gasseri* ATCC 33323, with emphasis on predicted functions that are likely to support the autochthonous nature of the organism in the human GIT. In addition, the phenotypic characterization of selected *L. gasseri* strains was carried out in order to compare the sequenced neotype strain to other strains and assess intraspecies diversity.

MATERIALS AND METHODS

Genome sequencing. Draft-phase genome sequencing was performed at the Joint Genome Institute, Walnut Creek, CA, under project 2662179 (http://genome.jgi-psf.org/finished_microbes/lacga/lacga.info.html). The draft-phase assembly was finished by Fidelity Systems, Inc., Gaithersburg, MD, using direct genome sequencing (http://www.fidelitysystems.com/). The complete genome sequence of *L. gasseri* and its updated annotation are found at GenBank under accession number CP000413.

Bioinformatic analysis. The complete genome sequence was automatically annotated by an extended version of GAMOLA (1). The gene model was adopted from the previously published gene model (68). Sequence similarity analyses were performed with the gapped BlastP algorithm (5), utilizing the nonredundant database provided by the NCBI (ftp://ftp.ncbi.nih.gov/BLAST/db) and a custom database comprising the currently published and completed *Lactobacillus* ORFeomes. A functional classification was applied using the COG (clusters of orthologous groups of proteins) database (95). Protein motifs were determined by Hmmer (http://hmmer.wustl.edu/) (32) using PFAM HMM libraries, with global and local alignment models (http://pfam.wustl.edu/) (11) and TIGRfam libraries with global and local alignment models (http://www.tigr.org/TIGRFAMs/) (48). In addition, InterPro (http://www.ebi.ac.uk/interpro/) (71) and gene ontology (GO) information (http://www.geneontology.org/) (7) was deduced from Pfam and TIGRfam hits where appropriate and incorporated into the final data matrix. Structural information including determination of tRNAs (tRNAscan-SE) (67) and prediction of signal peptide cleavage sites (SignalP) (15), transmembrane domains (TMHMM2) (62), and terminator-like structures (TransTerm) (36) was subsequently added to form a comprehensive functional genome layout.

Genome visualizations were obtained by using Genewiz (79). Sequence analyses were performed by in-house-developed software solutions. Metabolic pathway mapping using the ORFeome of *L. gasseri* ATCC 33323 was performed by using the software-suite PathwayVoyager (2) and the KEGG (Kyoto Encyclopedia of Genes and Genomes) online database (http://www.genome.ad.jp/kegg/kegg2.html).

A BLAST heat map based on the nonredundant BLAST database provided by the NCBI was constructed using in-house-developed software solutions. Briefly, the organism distribution on a genus level was identified for each predicted open reading frame (ORF). Corresponding e-values were grouped into ranges according to a customized window size. Threshold levels were defined for minimum overall frequency, and a strain or genus filter was applied where appropriate. The data visualization was realized by using SigmaPlot v9.01 (Systat Software, Inc., San Jose, CA) and by converting results into the long-form mesh data format.

Strains and culture media. The bacterial strains used in the present study are listed in Table 1. Strains were propagated statically at 37°C in MRS broth (Difco Laboratories, Inc., Detroit, MI) or on MRS agar supplemented with 1.5% agar. Carbohydrate utilization analyses were performed by using the API50 CH tests (bioMérieux, Durham, NC) according to the manufacturer's instructions.

Tissue culture. Caco-2 (ATCC HTB-37; American Type Culture Collection) epithelial cells were used between the 40th and 60th passages. Tissue cultures were maintained as described before by Buck et al. (20). All reagents used in maintenance of Caco-2 cells were obtained from Gibco (Gibco-Invitrogen Corp., Carlsbad, CA). Cells were grown on 15-mm Thermanox plastic coverslips (Nalge Nunc International, Rochester, NY) in treated Costar 12-well tissue culture plates (Corning, Inc., Acton, MA).

Adhesion assays. The adhesion of *Lactobacillus* strains was examined by using a modification of previously described methods (20, 25). Briefly, cultures were grown in MRS from a 1% inoculum of an overnight culture. After 16 h, 10-ml aliquots were centrifuged, washed with fresh MRS, and resuspended in 6-ml of MRS medium. The optical density of the cultures at 590 nm (OD_{590}) was adjusted to 0.50 to 0.54, and then the cultures were diluted 1:1 with the same medium. Portions (200 μl) of cells were applied to each plate well containing confluent Caco-2 cells, followed by incubation for 1.5 h at 37°C. After incubation, the monolayers were washed three times with phosphate-buffered saline (PBS), fixed with 1 ml of methanol, and Gram stained. Adherent bacterial cells were enumerated microscopically by examining 10 fields chosen randomly and averaging the results. Experiments were carried out in duplicate.

Growth and survival in bile. The growth of *Lactobacillus* strains was evaluated in MRS and MRS supplemented with 0.15, 0.25, 0.5, or 1% Oxgall (wt/vol; Difco) and automatically monitored by determining the changes in absorbance (A_{600}) as a function of time using a FLUOStar OPTIMA microtiter plate reader (BMG Labtech GmbH, Offenburg, Germany). The maximum specific growth rate was calculated from the slope of a linear regression line during exponential growth with a correlation coefficient (r^2) of 0.99. Each point represents the mean of three independent cultures.

Survival of early-log-phase (OD_{600} = 0.2 to 0.3) *Lactobacillus* cultures was examined in 7 and 10% Oxgall at pH 6 and pH 7. Cells were pelleted by centrifugation and resuspended in MRS broth (pH 6 or pH 7) or MRS broth supplemented with Oxgall. Cultures were held at 37°C for 10 min and then serially diluted and plated onto MRS agar by using a Whitley Automatic Spiral Plater (Don Whitley Scientific, Ltd., West Yorkshire, England). Survival was expressed as a ratio of survivors in Oxgall to survivors in MRS.

Survival in simulated gastric juice. Cells were grown overnight from a 1% inoculum in MRS. Aliquots (1 ml) were centrifuged, and the cells were washed three times with PBS (pH 7). The final pellet was resuspended in 1 ml of PBS, and 0.5-ml portions of cells were mixed with 2.5 ml of simulated gastric juice (24).

TABLE 1. Bacterial strains and primers

Lactobacillus strain or primer[a]	Characteristics or sequence (5′→3′)	Source or reference
Strains		
L. gasseri ATCC 33323	Human isolate, type strain	ATCC[b]
L. gasseri ADH	Lysogen strain	59
L. gasseri FR2	Isolated from endoscopy from healthy individual	63
L. gasseri JK12	Isolated from fecal sample of a healthy individual	63
L. gasseri ML3	Isolated from endoscopy from healthy individual	63
L. gasseri RF14	Isolated from fecal sample from healthy individual	63
L. gasseri RF81	Isolated from fecal sample from healthy individual	63
L. gasseri SD10	Isolated from fecal sample of healthy individual	63
L. gasseri WD19	Patient endoscopy isolate, bacteriocin producer	63
L. acidophilus NCFM	Human intestinal isolate	10
L. johnsonii 88*	Bacteriocin producer	54
L. helveticus ATCC 15009*	Type strain	ATCC
Primers (location in genome [bp])		
LGAS_0572 (600300)	ttgttgctccgcttgtaaatgtgtc	This study
LGAS_0573 (601026)	ggatctaagggcttaacagaatgtc	This study
LGAS_0634 (639805)	ggttaaatgttgacggaaacatcag	This study
LGAS_0636 (641058)	gatgatttccaccctcattctttac	This study
LGAS_0697 (680001)	caaggtatgacttggcatgaagaac	This study
LGAS_0699 (681281)	aacaccaagggagccatcttaatag	This study

[a] *, the phenotypic traits analyzed for *L. gasseri* and *L. acidophilus* strains were resistance to bile and simulated gastric juice, carbohydrate utilization, adhesion to Caco-2 cells, and oxalate degradation. *L. johnsonii* and *L. helveticus* were used as negative controls in oxalate degradation experiments.
[b] ATCC, American Type Culture Collection.

Aliquots (100 μl) were removed at 0, 60, and 120 min and enumerated on MRS plates in duplicate.

Degradation of oxalate. Decrease in oxalate concentration in culture supernatants was determined as previously described (8) with minor modifications. *Lactobacillus* strains were transferred two to three times in MRS containing 0.07 mM ammonium oxalate. Cells were then inoculated into the same medium, grown to an OD_{600} of 0.6, centrifuged, and resuspended in MRS containing 7.0 mM ammonium oxalate. Oxalate concentrations in the supernatants were measured in triplicate by using the Oxalate kit (Trinity Biotech, County Wicklow, Ireland).

PFGE. Preparation of genomic DNA samples and pulsed-field gel electrophoresis (PFGE) analysis were carried out as previously described (27, 94).

RESULTS AND DISCUSSION

General features and sequence comparison. The general features of *L. gasseri* ATCC 33323 are presented in Table 2 and Fig. 1. The genome is 1,894,360 bp in size, does not present plasmids, and is predicted to encode 1,810 genes, accounting for 88% of the sequence. The G+C content of the genome is 35.3%, which is similar to the G+C content of its closest relatives, *L. johnsonii* NCC 533 (34%) and *L. acidophilus* NCFM (34%), and lower than all of the eight completed *Lactobacillus* genomes (38 to 49%), with the exception of *L. salivarius* UCC 118 (32%). Thirteen putative transposon or transposon-related genes were identified in the genome sequence of *L. gasseri* (see Table S1 in the supplemental material). This number is close to those for *L. acidophilus* (19) and *L. johnsonii* (18) but significantly lower than that for *L. helveticus* (n = 225 [22]). *L. gasseri* contains 75 tRNA genes, in accordance with the number of tRNA genes and genome size of *Firmicutes* (97).

TABLE 2. *L. gasseri* ATCC 33323 genome features

Feature	Value
Genome size (no. of nucleotides)	1,894,360
No. of predicted ORFs	1,810
Avg ORF length in nucleotides	933
Coding percent	89.2
% G+C content	35.26
No. of TIGRfam domains	2,003
No. of PFam domains	2,107
No. of COG matches	1,540
No. of transmembrane domains	486
No. of tRNAs	75
No. of signal peptidase cleavage sites	271
No. of terminator structures	641

To investigate the broader phylogenetic position of *L. gasseri* to other genera, a BLAST heat map, excluding self-hits and based on the nonredundant BLAST database provided by the NCBI, was constructed by using in-house-developed software algorithms. The organism distribution on a genus level was identified for each predicted ORF and corresponding e-values were grouped into predefined ranges (Fig. 2). Not surprisingly, the most dominant heat flare was observed compared to other *Lactobacillus* species. Interestingly, the genus *Streptococcus* revealed a second, dominant flare within the *Lactobacillales*, partially reflecting the large number of sequenced and draft-phase genomes of this genus (41 at the time of analysis versus 3 for *Lactococcus*). When the heat map was restricted to one hit per genus (Fig. 2, bottom panel), a slightly more balanced distribution was observed, with the genera *Streptococcus*, *Lactococcus*, *Pediococcus*, *Enterococcus*, *Bacillus*, *Listeria*, and *Clostridium* revealing a similar distribution pattern. It is noteworthy that only very few hits were found between e-values of 1e-190 and 1e-150, possibly indicating an ancient diversification and/or substantial gene loss/gain during evolution, as indicated by Makarova et al. (68). Ancient divergence would yield significant levels of accumulated amino acid changes, resulting in reduced levels of similarities, whereas repeated cycles of gene acquisition and loss prior to further divergence

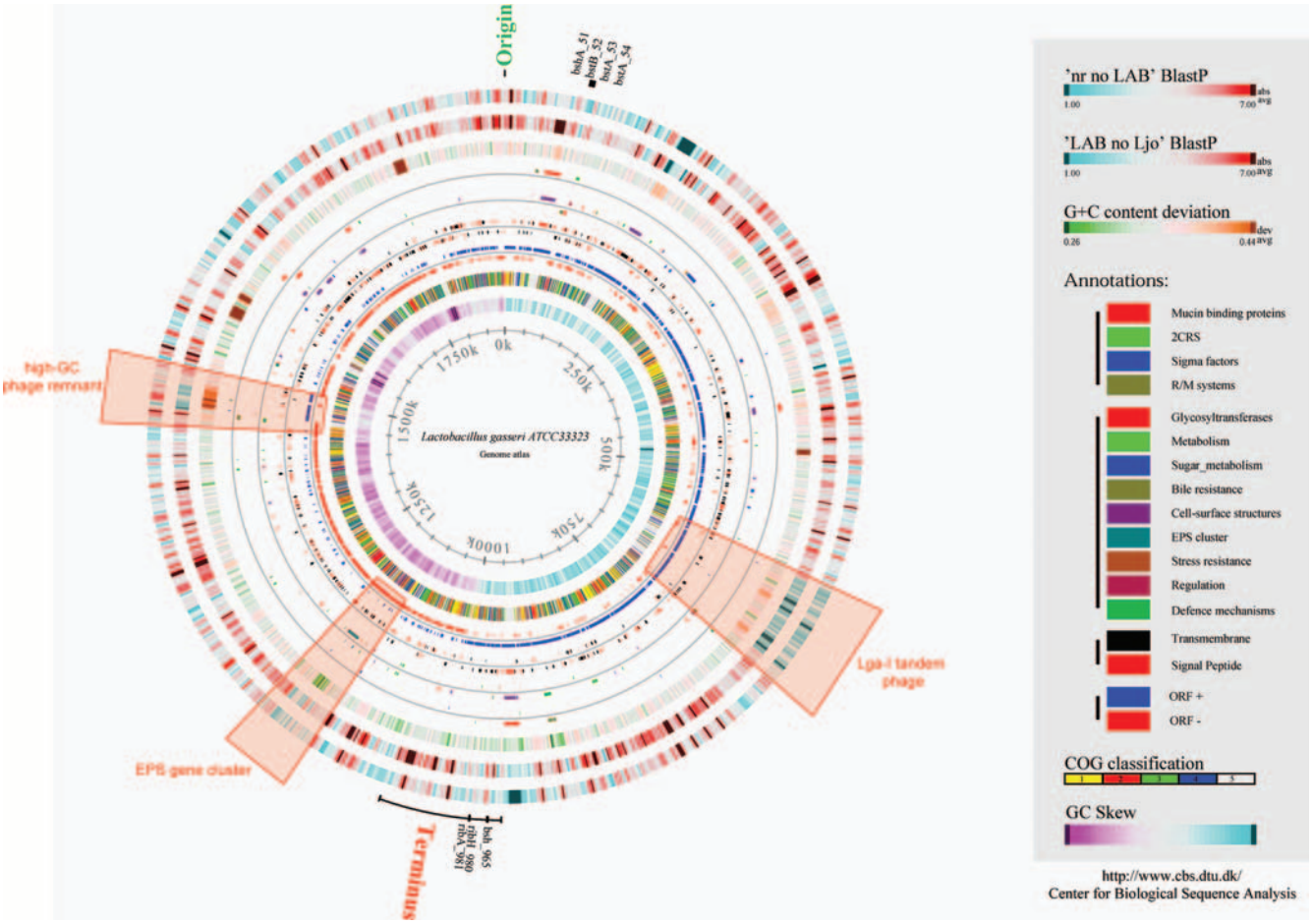


FIG. 1. Genome atlas of *L. gasseri* ATCC 33323. The circle was created by using Genewiz (79) and in-house-developed software. The right-hand legend describes the single circles in the top-down outermost-innermost direction. Outermost first ring, gapped BlastP (6) results using the nonredundant database minus published lactic acid bacterial sequences; second ring, gapped BlastP results using a custom lactic acid bacteria database, excluding the highly similar *L. johnsonii* NCC533 (83) genome. In both rings, regions in blue represent unique proteins in *L. gasseri*, whereas highly conserved features are shown in red. The degree of color saturation corresponds to the level of similarity. Third ring, G+C content deviation (green shading highlights low-GC regions, orange shading high-GC islands). Annotation rings 4 to 6, black vertical lines in the right-hand legend indicate ring-specific annotation grouping. Seventh ring, ORF orientation. ORFs in the sense orientation (ORF+) are shown in blue; ORFs oriented in the antisense direction (ORF−) are shown in red. Eighth ring, COG classification. COG families were assembled into five major groups: 1, information storage and processing; 2, cellular processes and signaling; 3, metabolism; 4, poorly characterized; and 5, ORFs with uncharacterized COGs or no COG assignment. Innermost ninth ring, GC-skew. Selected features representing single ORFs are shown outside of circle 1, with bars indicating their absolute size. The origin and terminus of DNA replication are identified in green and red, respectively. Three large genome islands harboring distinct features (EPS gene cluster and prophage makeup) have been highlighted with red trapezoids.

would result in highly similar sequence levels. Also, the presumed ancient split from *Bacillales* would likely leave remnants of a conserved core genome, further contributing to the number of highly conserved sequences (e-values below 1e-150). Notably, within the phylum *Firmicutes*, only the genera *Bacillus*, *Listeria*, and *Clostridium* showed significant levels of similarities, rivaling those of *Lactobacillales*. An initial high-frequency core above e-values of 1e-80 and an isolated zone of highly conserved sequences (e-values below 1e-150) mirrored the *Bacillus* flare structure. Given the remarkable ability of *Bacillus*, *Listeria*, and *Clostridium* to survive in various environmental niches such as those co-occupied with lactobacilli, this observation might indicate an extensive level of directional gene transfer. Further analyses are required to investigate this close relationship between these two genera and whether the exchange of genetic information is uni- or bidirectional.

Genera outside the phylum *Firmicutes* revealed only low-level similarities to *L. gasseri*, predominantly to *Pseudomonas*, *Shewanella*, *Vibrio*, *Geobacter*, and *Burkholderia* (not exceeding evalues of 1e-30), suggesting strong barriers toward interphylum horizontal gene transfer (HGT).

General metabolic pathways. Overall, genome similarity is highly significant ($E \leq$ 1e-100) between *L. johnsonii* NCC 533 and *L. gasseri* ATCC 33323 given that ca. 50% of the predicted ORFs in *L. gasseri* show similarity to genes in *L. johnsonii* (58, 83). Consequently, these species share a number of metabolic capabilities. *L. gasseri* shows a partial citrate cycle with a putative operon (LGAS_0850 to LGAS_0853) containing a fumarate reductase (*frdA*), fumarate hydratase (*fum*C), a malate

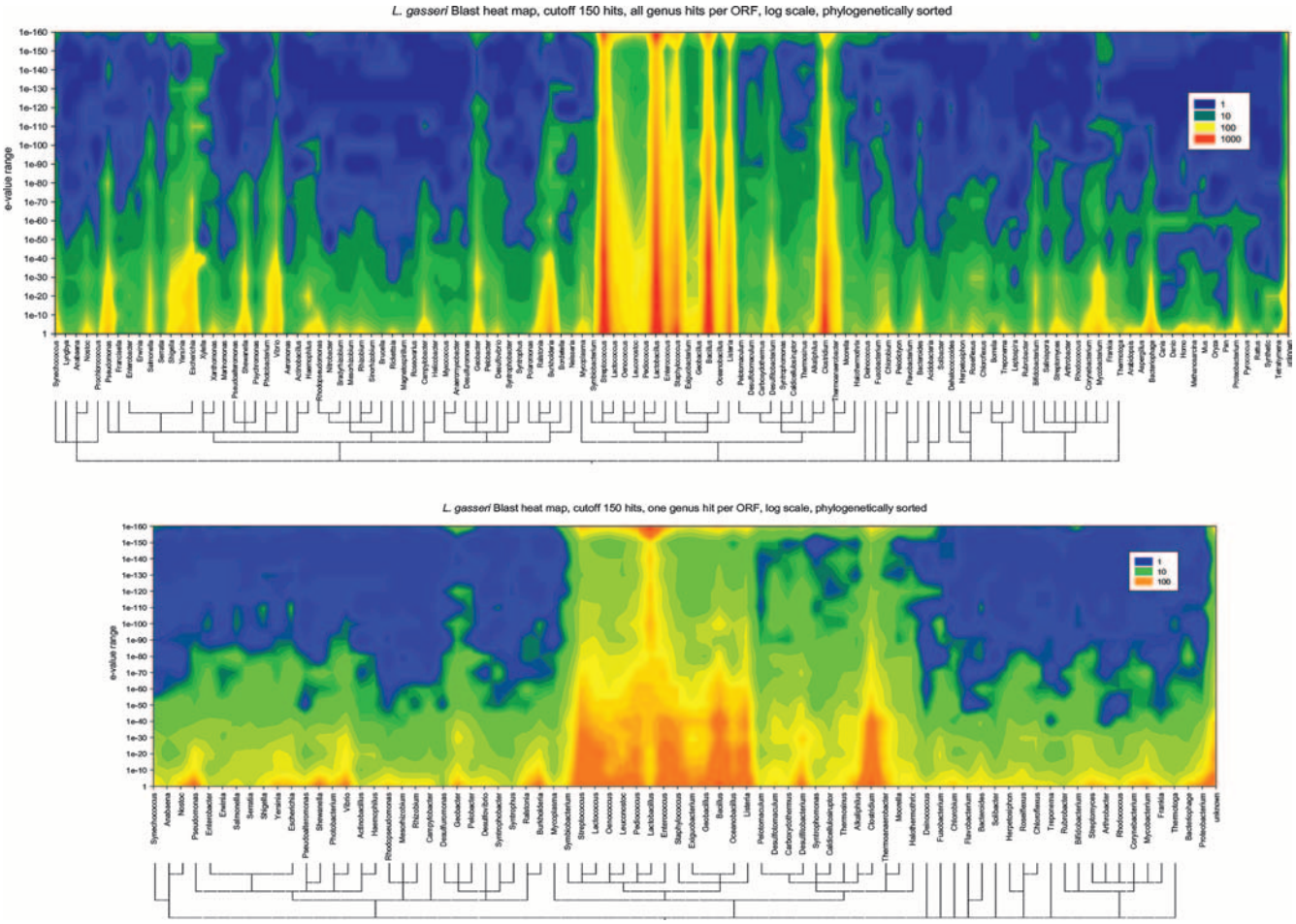


FIG. 2. BLAST result distribution across the *L. gasseri* ATCC 33323 ORFeome. In both figures, the *x* axis (horizontal axis) shows all genera with at least 150 BLAST hits throughout the ORFeome. Genera are phylogenetically sorted based on a semidynamically reparsed phylogenetic tree obtained from the Ribosomal Database Project II (RDPII) (http://rdp.cme.msu.edu/hierarchy/hierarchy_browser.jsp), selecting NCBI taxonomy, level 10 genera display list, and set to include archaeal sequences. Bacterial or archaeal genera not covered within the RDPII data are entered and parsed from a separate data file, when appropriate. Phylogenetic distribution and grouping of genera is indicated using an ASCII based tree-abstraction. The *y* axis indicates the e-value ranges, and the *z* axis (color coded) represents the frequency of hits for each genus in each e-value range in log scale. Respective log color scales are indicated in each figure, whereby warmer colors indicate higher frequencies. The bottom panel uses a frequency cutoff of one hit per genus per ORF, effectively limiting the hit rate to the best BLAST hit found in each given ORF and genus. The top panel allows all BLAST hits per genus per ORF, accepting multiple genus hits per ORF.

dehydrogenase (*mdh*), and a conserved unknown protein. The generated oxaloacetate could be then converted into phosphoenolpyruvate by the phosphoenolpyruvate carboxykinase (EC 4.1.1.49, LGAS_0149) or subsequently used in the biosynthesis of aspartate through an aspartate aminotransferase (LGAS_1143).

Amino acid metabolism. In addition to the ability to synthesize aspartate from oxaloacetate, *L. gasseri* putatively dedicates three enzymes (LGAS_0854, *ansA*; LGAS_1760, *asnA*; and LGAS_0133, *asnB*) to the conversion of L-aspartate into L-asparagine. As in *L. acidophilus* NCFM (4), *L. gasseri* is also able to convert L-aspartate into L-homoserine and finally threonine through a cascade of five reactions (LGAS_1089, *thrA2*; LGAS_1088, *asd*; LGAS_1087, *thrC*; LGAS_1086, *thrA*; and LGAS_1085, *thrB*). A putative threonine dehydratase (EC 4.3.1.19, LGAS_1436) could use L-threonine as a substrate to produce 2-oxobutanoate and ammonia. As in *L. johnsonii*, the pathway implicated in the generation of L-isoleucine, L-leucine and L-valine from 2-oxobutanoate, does not appear to be present in *L. gasseri*.

Nonessential amino acids have a central metabolic role in the human intestine, as was specifically demonstrated for glutamine, and its precursor, glutamate. In fact, there is evidence that glutamine, especially systemic glutamine, supports the function of the intestinal mucosal system (84). In addition, glutamate is a key intermediate in the metabolism of amino acids in lactic acid bacteria since aminotransferases use this amino acid as the donor substrate of amino groups (40). In *L. gasseri*, a glutamic-oxaloacetic transaminase (EC 2.6.1.1, LGAS_1143) might act on L-aspartate and 2-oxoglutarate to generate L-glutamate, which could then be converted to D-glutamate through a glutamate racemase (LGAS_0426) and to L-glutamine through glutamine synthetase (LGAS_1392) and glutaminase (LGAS_0504).

Like *L. acidophilus*, *L. gasseri* has the potential to generate L-cysteine from pyruvate and L-methionine from L-cysteine through the intermediates cystathionine and L-homocysteine. However, unlike *L. acidophilus*, the metabolic pathway necessary to convert L-aspartate into L-lysine does not appear to be complete in *L. gasseri*. Serine could also be directly generated from pyruvate. In addition, *L. gasseri* shows an almost complete pathway that could generate serine from glycerate through four enzymatic steps. A serine hydroxymethyltransferase (*glyA*, LGAS_0254) could potentially act on serine to generate glycine.

Purine and pyrimidine metabolism. The enzymes required to generate 5-phosphoribosyl-1-pyrophosphate (EC 3.6.1.13, LGAS_1211; EC 2.7.6.1, LGAS_0097 and LGAS_0208) are present in *L. gasseri*. In addition, some similarity was observed between ORF LGAS_1763 and a putative amidophosphoryltransferase (e-value 3e-23) from *Chromobacterium violaceum*, necessary to generate 5-phosphoribosyl amine. Six of the subsequent nine enzymes required to generate IMP seem to be absent in *L. gasseri*. However, ATCC 33323 could generate GMP, xanthosine 5′-phosphate, and AMP from IMP. A putative 5′ nucleotidase (EC 3.1.3.5, LGAS_0440) could produce the corresponding nucleoside, substrates for the predicted nucleoside hydrolases LGAS_0874 and LGAS_0559. The pyrimidine requirements of cells can be satisfied either via de novo synthesis or via salvage of preformed pyrimidine bases and nucleosides provided by the surrounding medium. All of the necessary genes for de novo synthesis of purine samples appear to be present in *L. gasseri* ATCC 33323.

Metabolism of cofactors. The contribution of intestinal bacteria to the biosynthesis of vitamins and cofactors in the GIT was recognized as early as 1942 (21). Specifically, the presence of a complete pathway for riboflavin synthesis, and the ability to produce and excrete small amounts of the compound has been reported in *L. lactis* (52). Unlike *L. johnsonii* and *L. acidophilus*, *L. gasseri* presents a more complete, but still partial, pathway for the synthesis of riboflavin. A potential operon, not present in *L. johnsonii* or *L. acidophilus* (LGAS_0981 and LGAS_0980), contains the GTP cyclohydrolase II (*ribA*, EC 3.5.4.25) and riboflavin synthase (*ribH*, EC 2.5.1.9) enzymes required for the first and last steps in the synthesis of riboflavin from GTP. Only one enzyme, the 5-amino-6-(5-phosphoribosylamino) uracil reductase RibD, appears to be absent in *L. gasseri*. In addition, *L. gasseri* appears to be capable of converting riboflavin into flavin mononucleotide via riboflavin kinase (EC 2.7.1.26, LGAS_0820) and flavin mononucleotide to riboflavin via a putative phosphotyrosine protein phosphatase (EC 3.1.3.2, LGAS_1180).

Like most lactobacilli (41), *L. gasseri* appears to be incapable of de novo synthesis of NAD. The presence of two key enzymes, nicotinamidase (EC 3.5.1.19, LGAS_1097) and nicotinate phosphoribosyltransferase (NAPRTase, EC 2.4.2.11, LGAS_1542), indicates that this organism can utilize both nicotinate and nicotinamide to generate NAD. Similar to *L. johnsonii*, *L. gasseri* seems to be deficient for the enzymes required to synthesize thiamine, biotin, pyridoxine, and panthothenate. The specific sodium:panthothenate transporter could not be identified in *L. gasseri* by sequence analysis, but the necessary genes for the synthesis of coenzyme A from panthothenate are present.

Annotation for functions important in the GIT. (i) Sugar transport and metabolism. The carbohydrate utilization repertoire of *L. gasseri* resembles that of *L. johnsonii* (83), where the majority of the sugar transporters are devoted to the uptake of monosaccharides, disaccharides and, to a lesser extent, trisaccharides. The *L. gasseri* genome encodes 21 putative phosphoenolpyruvate sugar phosphotransferase systems (PTS), with predicted specificities for fructose, mannose, glucose, cellobiose, lactose, sucrose, trehalose, β-glucosides, and *N*-acetylglucosamine (GlcNAc) (see Table S2 in the supplemental material). In addition to PTSs, three putative sugar ATP-binding cassette (ABC)-type transport systems were identified with predicted specificities for maltose (LGAS_0214 to LGAS_0217), galactosides and/or pentoses (LGAS_0256 to LGAS_0258), and the third one with unknown specificity (LGAS_1607 to LGAS_1609). Genes coding for two glucose uptake permeases were also present (LGAS_0170, LGAS_1051), one of which may also transport ribose. Unlike the closely related species *L. acidophilus* and *L. johnsonii*, *L. gasseri* does not encode a lactose/galactose permease or a β-galactosidase. Instead, as previously noted by Pridmore et al. in *L. johnsonii* (83), lactose uptake and intracellular hydrolysis are likely mediated by two PTS transporters (LGAS_0339-LGAS_0340, LGAS_0497-LGAS_0498) and four putative phospho-β-galactosidases (LGAS_0182, LGAS_0341, LGAS_0499, and LGAS_0638), respectively.

Interestingly, of seven putative cellobiose PTS transporters identified in *L. gasseri*, only one has all of the structural components IIA (EIIA), B (EIIB), and C (EIIC) (encoded by LGAS_0189, LGAS_0191, and LGAS_0192, respectively) (see Table S2 in the supplemental material). The remaining six cellobiose transporters consisted of only the substrate-specific EIIC permease of a PTS system. A recent study showed that the expression of similar "orphan" cellobiose PTS EIIC components was induced in *L. plantarum* during passage through the mouse GIT model system (18). One of these cellobiose PTS EIIC proteins, encoded by LGAS_0646, has no homolog among the lactic acid bacteria. Rather, this transporter showed 62 to 63% sequence identity to the PTS EIIC components from *Enterococcus faecium* (GenBank accession no. ZP_00604348) and *E. faecalis* (GenBank accession no. NP_815915), suggesting lateral transfer of sugar utilization genes among the cecal microflora.

The ATCC 33323 genome encodes 20 putative glycosyl hydrolases (Fig. 1), mostly glucosidases and galactosidases, with predicted substrate specificities for a diversity of di- and trisaccharides. Two of these sugar hydrolases, a β-glucosidase (LGAS_0394) and a putative glycosyl hydrolase family 31 (LGAS_0517), have no homolog among the lactic acid bacteria. The former shared moderate sequence identity (~40%) to the corresponding hydrolases from species of *Bacillus*, *Clostridium*, and *Listeria*, whereas the latter was found most closely related to glycosyl hydrolases from *Clostridium* and *Salmonella* (~50% identity).

A putative neopullulanase or maltogenic amylase (LGAS_0211) was found encoded within the maltose/maltotriose utilization gene cluster (LGAS_0210 to LGAS_0218), which may potentially degrade pullulan, a linear polysaccharide consisting of maltrotriose units linked by α-1,6-glucosidic bonds. The enzyme shared 75 to 90% sequence identity to neopullulanases in *L. johnsonii* (accession no. NP_964228) and *L. acidophilus* (acces-

sion no. YP_194702). A recent study showed that *L. gasseri* ATCC 33323 was unable to utilize pullulan as a sole carbon source (S. O'Flaherty and T. R. Klaenhammer, unpublished data). In addition, no amylolytic activity toward β-cyclodextrin was detected in culture supernatant or crude cell extract of ATCC 33323 (76). However, the purified neopullulanase/maltogenic amylase was able to hydrolyze pullulan, β-cyclodextrin, and soluble starch (76). Nevertheless, this enzyme lacks a signal peptide sequence, indicating its cytoplasmic activity toward lesser complex physiological substrates, such as maltotriose, that are more likely to be accumulated by the encoded sugar transporters.

To evaluate the diversity among *L. gasseri* strains and to further investigate the type of carbohydrate sources that are able to support the growth of *L. gasseri* ATCC 33323, the sugar utilization profile of the strain was assessed along with a group of eight human *L. gasseri* isolates from our culture collection by using the API50 CH test. The results summarized in Table S3 in the supplemental material showed that all *L. gasseri* strains were able to ferment glucose, fructose, cellobiose, trehalose, sucrose, GlcNAc, and esculin. Other hexoses and disaccharides, including mannose, galactose, tagatose, gentibiose, and maltose, and the modified β-glucosides amygdaline, arbutin, and salicin, were fermented at variable degrees among the strains. One of the *L. gasseri* strains, WD19 (originally isolated from a patient endoscopy), showed a similar sugar fermentation pattern to that of strain ATCC 33323, including its ability to utilize lactose and starch that were nonfermentable by the other *L. gasseri* strains. It remains to be determined whether the capability of *L. gasseri* ATCC 33323 to utilize starch as a carbon source is encoded by the maltose/maltotriose utilization gene cluster, particularly the involvement of the putative neopullulanase/maltogenic amylase as discussed previously.

The API50 CH analysis also revealed GlcNAc as a readily fermentable hexosamine among the *L. gasseri* strains. This aminosugar, along with *N*-acetyl-D-galactosamine, D-galactose, sialic acid, and L-fucose, is a component of the oligosaccharide side chains of epithelial mucin glycoproteins (74). Thus, secreted mucin glycoproteins that pass through or present in the GIT represent another important nutrient source for the resident microflora. Although a limited number of intestinal species has been proved to be able to hydrolyze complex carbohydrates derived from gastrointestinal mucins (e.g., *Ruminococcus torques* and some strains of *Bifidobacterium* [51]), a clear enrichment for genes involved in mucin degradation has been revealed in the ongoing microbiome metagenomics project (101). Our study showed that supplementation of porcine gastric mucin as a sole carbon source did not support the growth of *L. gasseri* ATCC 33323 or any of the other *L. gasseri* strains (data not shown). In addition, enzymes that are required to degrade mucin, including neuraminidase, α-fucosidase, α-*N*-acetylgalactosaminidase, β-*N*-acetylglucosaminidase, or β-galactosidase, did not appear to be present in the genome of ATCC 33323. On the other hand, ATCC 33323 possesses at least one PTS transporter that likely participates in the cross-feeding of free GlcNAc moieties possibly derived from mucin degradation by other intestinal microbes. The resulting intracellular *N*-acetylglucosamine-6-phosphate would serve as a substrate for the metabolism of fructose, mannose, and amino acids, as well as for generating precursors that feed into the peptidoglycan biosynthetic pathway. Close examination of the aminosugar metabolic pathway of *L. gasseri* ATCC 33323 indicated the presence of an *N*-acetylmannosamine-6-phosphate epimerase, NanE (encoded by LGAS_1658), which is responsible for the conversion of *N*-acetylmannosamine-6-phosphate to *N*-acetylglucosamine-6-phosphate in the sialic acid utilization pathway. However, the absence of genes encoding an *N*-acetylneuraminate lyase and a *N*-acetylmannosamine kinase involved in the initial steps of sialic acid catabolism rules out the possibility of sialic acid being used as an energy source by *L. gasseri*.

None of the pentoses, sugar alcohols, or deoxysugars in the API50 CH studies were fermented by any of the *L. gasseri* strains or *L. acidophilus* NCFM. Although putative transporters were predicted for the uptake of ribose (see Table S2 in the supplemental material), *L. gasseri* was unable to use ribose as a carbon source. The apparent lack of genes encoding a transketolase and a transaldolase in the pentose phosphate pathway of ATCC 33323, the latter enzyme of which is also missing in *L. acidophilus*, rendered these strains incapable of assimilating ribose. Finally, no fermentation activity of the oligosaccharides melezitose or raffinose was observed among the *L. gasseri* strains. Unlike *L. acidophilus* NCFM, none of the *L. gasseri* strains were able to grow on the FFn-type fructooligosaccharides, a widely used prebiotics, as a sole carbon source. Furthermore, a recent study by Ward et al. (103) showed that *L. gasseri* ATCC 33323 was not able to ferment breast milk oligosaccharides. On the other hand, Martin et al. (69) investigated the presence of LAB in the breast milk of healthy women and identified most of the isolates as *L. gasseri*.

Despite the absence of β-fructosidase, inulinase, xylanase, arabinosidase, or other glycosyl hydrolases that are required for the utilization of complex carbohydrates in *L. gasseri*, its diversified PTS transporters and sugar hydrolases for mono- and disaccharides imply its competitive fitness in the upper gastrointestinal environment, where these readily fermentable sugars are present. In *L. plantarum*, the expression of sugar PTS for sucrose, cellobiose, and *N*-acetylglucosamine/galactosamine, as well as several sugar hydrolases, was induced in situ in the GIT (18). Similarly, in vivo induction of genes involved in the metabolism of β-glucosides were observed in *Listeria monocytogenes* and *Streptococcus gordonii* (42, 55). These studies signify the importance of the noncomplex sugar utilization genes that not only fulfill energy requirements but may also be involved in other physiological functions that contribute to the survival and persistence of *L. gasseri* in the gut ecosystem.

(ii) BSH, bile transporters, and drug resistance. For enteric species, bile is a major component encountered in the GIT. Some species, including many probiotic bacteria, possess the ability to interact with and modify bile salts by hydrolyzing the amide linkage between their amino acid moieties and cholesterol backbones via bile salt hydrolases (BSH) (13). *L. gasseri* contains similarity (LGAS_0051 to LGAS_0054) to a locus in *L. acidophilus* KS-13 and *L. johnsonii* 100-100 encoding a BSH (LGAS_0051; BshA) (Table 3) (34). This region of similarity is syntenous with the sequenced genome of *L. johnsonii* NCC 533. In *L. gasseri*, the *bshA* gene is present in a putative operon with two transporters of the major facilitator superfamily, *bstA* and *bstB*. Putative homologues of these genes in *L. johnsonii* 100-100 were shown to play a role in bile salt transport (34).

TABLE 3. Highest BLAST hits for *L. gasseri* BSH

Bsh type and species	% Identity	% Positive
BshA (LGA0051)		
L. acidophilus	98	99
L. johnsonii	93	97
L. johnsonii NCC533 (LJ0056)	93	96
L. acidophilus	93	96
B. adolescentis	43	61
Bsh (LGA0965)		
L. johnsonii NCC533 (LJ1147)	66	80
L. acidophilus NCFM (LBA1078)	65	77
L. acidophilus NCFM (LBA0892)	57	73
L. johnsonii NCC533 (LJ1412)	56	74
L. johnsonii	56	74
L. reuteri 100-23	55	74
L. monocytogenes EGD-e (LMO2067)	50	66

Interestingly, there are two copies of one transporter gene, *bstA*, in the *L. gasseri* genome. One copy appears to encode only 137 amino acids of the N-terminal region of the protein, while upstream, the entire 252-amino-acid protein is present. The truncation in this locus is also evident in *L. johnsonii* NCC 533. Another BSH, LGAS_0965 shows a high degree of homology with the BSH in *L. johnsonii* NCC 533 and *L. acidophilus*. The homologue of this gene in *L. acidophilus* NCFM was shown to specifically hydrolyze bile salts conjugated to the amino acid taurine (70).

The detergent-like properties of bile can be damaging to bacterial cell membranes, proteins, and DNA (12); therefore, the ability to tolerate bile in the GIT is a major requirement for intestinal strains. Mechanisms of bile tolerance are not well understood, but increasing numbers of sequenced genomes coupled with functional genomic analysis are promoting their elucidation. *L. gasseri* contains homology (LGAS_0711 to LGA_0713) to part of an operon in *L. acidophilus* NCFM that has been functionally shown to play a role in bile tolerance (81). The proteins encoded by this putative operon in *L. gasseri* include a two-component regulatory system (2CRS) and a possible homologue of RelA, a protein that plays a role in gene regulation during times of stress. This putative operon also contains three hypothetical proteins not seen in *L. acidophilus*. Strains with mutations in the 2CRS and RelA homologue in *L. acidophilus* showed a significant decrease in recovery in the presence of bile (81). Immediately upstream of this operon is a locus (LGAS_0706-LGAS_0707) containing homology to BilEAB, an ABC transporter in *L. monocytogenes* that plays a role in both bile efflux and virulence in this strain (90).

Both the growth and the survival of *L. gasseri* were examined in the presence of bile. Figure 3 shows the reduction of the maximum specific growth rate (μ_{max}, h^{-1}) with increasing concentrations of Oxgall in the culture medium. The analyzed strains were subjectively grouped according to their sensitivity to this compound in "high" ($\mu_{max} = 0$ at 0.25% Oxgall, Fig. 3A), "medium" ($\mu_{max} = 0$ at 0.5% Oxgall, Fig. 3B), and "low" ($\mu_{max} > 0$ at 0.5% Oxgall, Fig. 3C) concentrations. There was high intraspecies variability, with ATCC 33323 located in the "medium" group. As observed in Fig. 3, the growth of all of the *L. gasseri* strains was reduced in MRS with 0.5% Oxgall, a concentration that had less of an effect on the growth of *L. acidophilus* NCFM (Fig. 3C).

The survival of these species was also examined in MRS with 7 and 10% Oxgall at pH 6 and pH 7 (Fig. 4). Previous work has shown the importance of pH in bile tolerance, with cell survival in bile decreasing with lowering pH (14). The survival of *L. gasseri* in bile at pH 7 is comparable to that of *L. acidophilus* NCFM, but the ratio of *L. gasseri* survivors is approximately an order of magnitude higher at pH 6. No differences in survival were observed between MRS at pH 6 and pH 7.

It has been suggested that one of bile's antimicrobial properties is the ability of the bile salts to acidify the cytoplasm of a cell by passing through the membrane in protonated, uncharged, forms (29). At lower pH, more bile salts are proto-

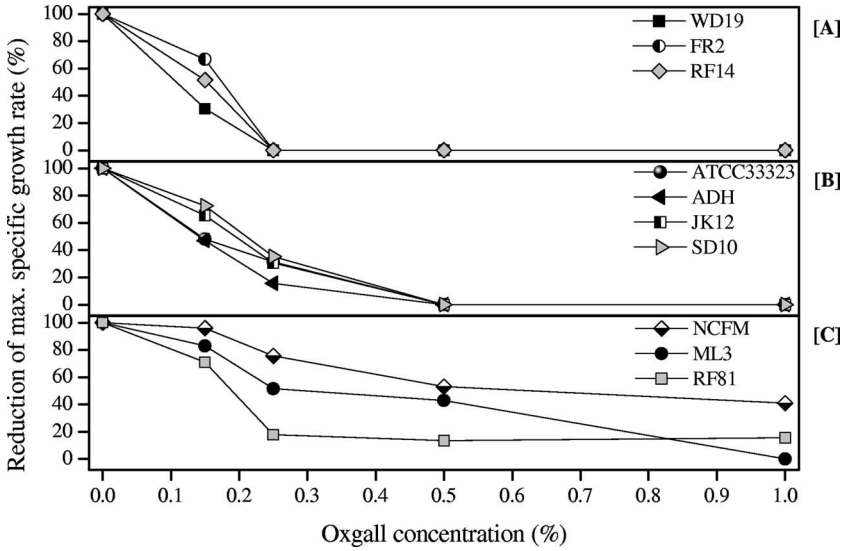


FIG. 3. Percent reduction of the maximum specific growth rate (μ_{max}) of *L. gasseri* strains and *L. acidophilus* NCFM exposed to increasing concentrations of Oxgall. (A) Strains highly sensitive to bile ($\mu_{max} = 0$ at 0.25% Oxgall); (B) strains of medium sensitivity ($\mu_{max} = 0$ at 0.5% Oxgall); (C) strains with the lowest sensitivity to bile ($\mu_{max} > 0$ at 0.5% Oxgall). Each point represents the mean of three replicates.

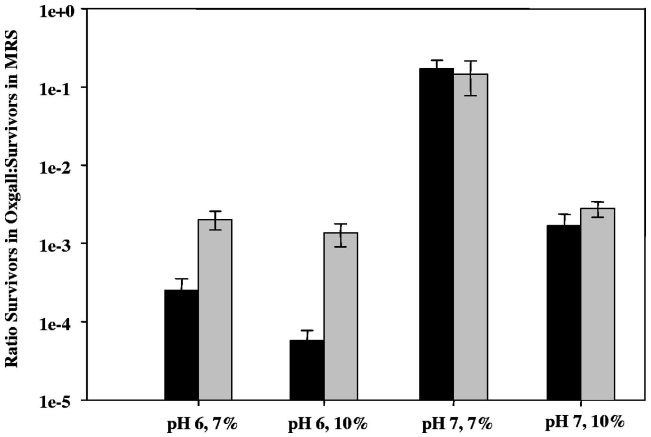


FIG. 4. Ratio of survival of early log phase *L. gasseri* ATCC 33323 (▨) and *L. acidophilus* NCFM (■) at designated concentrations of Oxgall to survival in MRS. Error bars represent the standard error of the mean for three replicates.

nated and thus are able to enter the cell, where the proton dissociates from the bile salt, causing cytoplasmic acidification. It is possible that the genetic repertoire of *L. gasseri* encodes proteins that make it better able to tolerate this type of stress over *L. acidophilus*, but the mechanism of this tolerance is unknown.

Multidrug transporters have received attention recently due to findings that they can play roles in the transport of substrates other than antibiotics, including bile and steroid hormones (82). Overall, the *L. gasseri* genome contained 19 multidrug transporters (see Table S4 in the supplemental material), including 2 putative bile salt transporters (LGAS_0052 and LGAS_0054). While most of these transporters contained extensive homology to proteins in *L. johnsonii* and *L. acidophilus*, LGAS_1423 was most similar to the cyanobacterial genus *Synechococcus*. This similarity was weak, however, with only 25% identity to this protein. LGAS_0976 contained similarity to a macrolide efflux protein in *E. faecium*, which could play a role in resistance to antibiotics such as erythromycin.

(iii) Cell surface structures. The cell surface of gram-positive bacteria contains a range of cell membrane- and cell wall-associated proteins and cell surface structures that directly affect the interactions of the bacteria with the host (17, 98). Analysis of the *L. gasseri* genome revealed 271 predicted proteins with a putative signal peptide sequence. A putative cleavage site was detected for 134 of these proteins through the action of two predicted signal peptidases I (LGAS_0793 and LGAS_1116). Most of these proteins are predicted to be anchored to the membrane via a transmembrane helix or a N-terminal lipoprotein domain (see Table S5 in the supplemental material), whereas only seven are anchored to the peptidoglycan via an LPxTG motif.

With 14 putative mucus-binding proteins, the genome of *L. gasseri* has the greatest number of proteins of this family among the sequenced lactobacilli (Fig. 1 and 5). Six of them show a signal peptide (LGAS_0041, LGAS_0945, LGAS_0946, LGAS_1623, LGAS_1632, and LGAS_1699), and four of those (LGAS_0041, LGAS_0945, LGAS_1632, and LGAS_1699) are attached to the membrane via transmembrane domains and covalently anchored to the cell wall via the C terminus, since their sequences exhibit an LPxTG motif putatively cleaved by a sortase A (LGAS_0828). These predicted proteins contain 6 to 12 copies of a conserved mucus-binding domain (Pfam PF06458) at the C-terminal end and share similarities ($1^{-127} < E < e^{-31}$) with the MUB protein of *L. reuteri* 1063, shown to be involved in adhesion to mucin (86), and with the *L. acidophilus* NCFM protein LBA1392 ($1^{-138} < E < 1^{-16}$) that mediates adhesion to Caco-2 cells (20). It is notable that eight additional ORFs were also identified that encode putative mucus-binding proteins, but these do not show a signal peptide (see Table S5 in the supplemental material). Among them, ORFs LGAS_0407, LGAS_1641, and LGAS_1624 appear to be truncated proteins, and resequencing of these regions confirmed the gene organization (data not shown). LGAS_0407 and LGAS_1641 also contain the conserved domain "Rib/alpha-like repeat" present in Rib and alpha surface antigens of group B streptococci (IPR012706) that have been shown to promote adhesion to human epithelial cells (92). The presence of other putative mucus-binding proteins that did not show a signal peptide but that are predicted to have large extracytoplasmic domains indicates that they might be nonfunctional or that they might be secreted via another mechanism. In our study, proteins with predicted putative transmembrane domains, but no signal peptide, were considered nonsecretory proteins.

A predicted lipoprotein manganese/zinc ABC transporter (LGAS_1696) may also act as an adhesin, based on the presence of two amino acid repeats (IPR006128 and IPR006129), and similarities with adhesins PsaA and SsaB from *Streptococcus pneumoniae* and *Streptococcus sanguis* (43, 88). The analysis of the *L. gasseri* genome also revealed the presence of FbpA, a putative fibronectin-binding protein (LGAS_1023) that may mediate adhesion of *L. gasseri* to fibronectin, a glycoprotein of the extracellular matrix of epithelial cells. Both proteins have homologs in the related species *L. johnsonii* NCC 533 and *L. acidophilus* NCFM.

Another interesting feature of *L. gasseri*'s potential cell surface structures is the presence of a putative exopolysaccharide (EPS) gene cassette (Fig. 1 and see Table S6 in the supplemental material). EPSs are carbohydrate polymers that can be secreted into the environment or remain attached to the cell wall. These polymers have been shown to influence adhesion properties of probiotic and pathogenic strains (87). In particular, heteropolysaccharides consist of structurally identical repeated subunits composed of at least two different monosaccharides linked by different types of glycosidic bonds. Analysis of the genome sequence of *L. gasseri* revealed the presence of an EPS cluster composed of 16 genes (LGAS_1156 to LGAS_1172). As already described for other lactic acid bacteria (reviewed by De Vuyst and Degeest [30]), the EPS region was organized into four functional regions: (i) a central region consisting of nine predicted glycosyl transferases (LGAS_1160 to LGAS_1169) specifically required for the addition of specific carbohydrates and the catalysis of specific glycosidic bonds; (ii) two regions flanking the central region encoding proteins putatively involved in polymerization and export (LGAS_1156 to LGAS_1159, and LGAS_1169 to LGAS_1171); and (iii) a regulatory region located at the beginning of the gene cluster (LGAS_1172). The unusually high number of

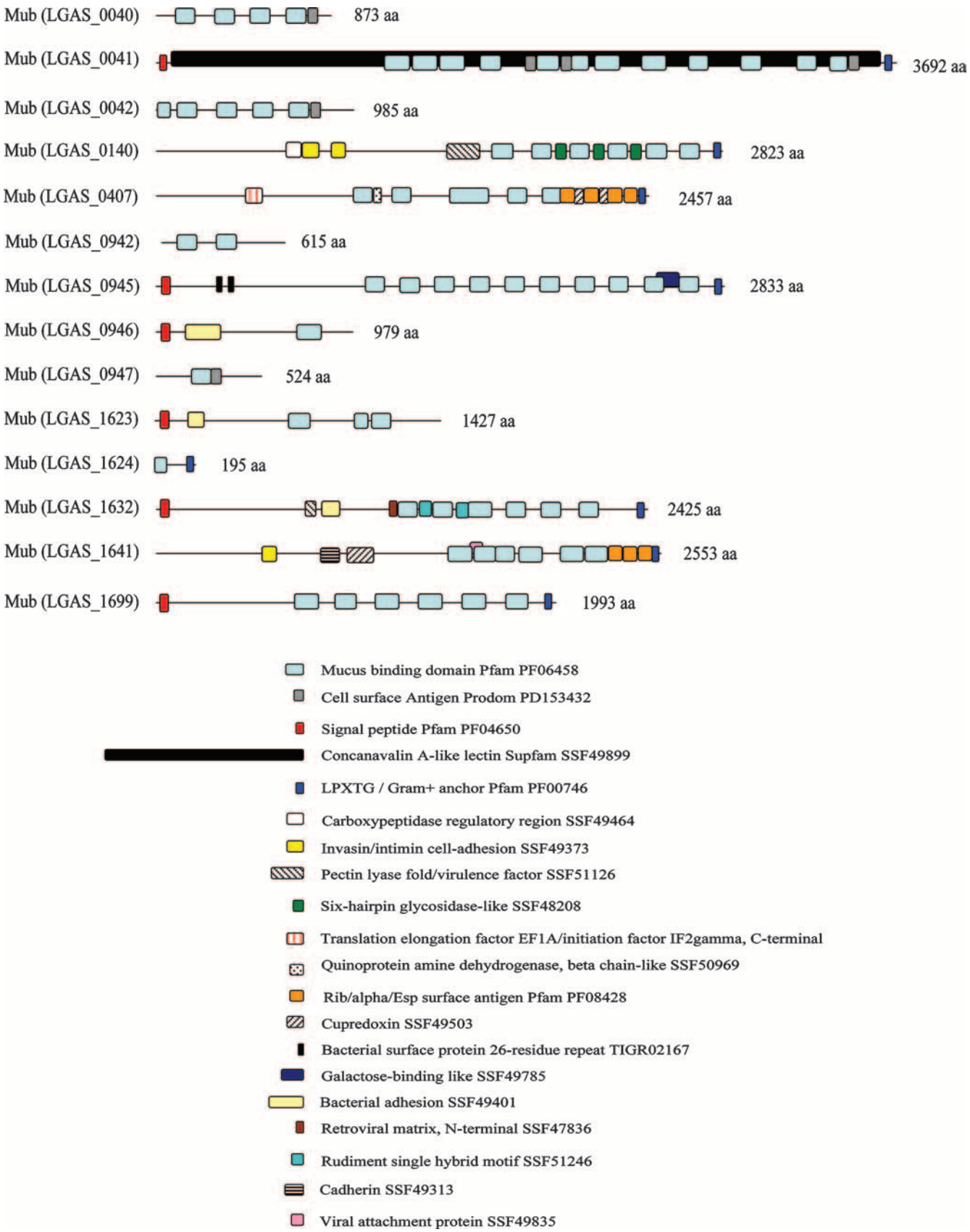


FIG. 5. Conserved domain analysis of *L. gasseri* ATCC 33323 putative mucus-binding proteins. Conserved domains according to Interproscan (http://www.ebi.ac.uk/InterProScan/) are indicated in the figure.

TABLE 4. Relative cell adhesion levels and affects of simulated gastric juice on the viability of selected *L. gasseri* strains and *L. acidophilus* NCFM strains

Strain	Adhesion[a]	% Survival in simulated gastric juice at pH 2.0 at:	
		60 min	120 min
L. gasseri			
ATCC 33323	29.4	23.7	0
ADH	30.1	75.1	0
FR2	36.3	95.7	0.4
JK12	34.6	6.6	0
ML3	13.6	100	20.6
RF14	13.2	14	0
RF81	14.9	100	62.2
SD10	26.3	60.3	0.4
WD19	15.1	61.2	27.5
L. acidophilus NCFM	54.7	1	0

[a] The values indicate the number of bacterial cells adhered to Caco-2 cells, enumerated microscopically. Each number is the mean of 10 counted fields chosen randomly.

glycosyltransferases (see Table S6 in the supplemental material) suggests that the potential synthesized polymer could be of a high complexity. As in *L. acidophilus* NCFM, the presence of two putative transposases (LGAS_1154 and LGAS_1155) downstream of the *eps* cluster and the low G+C content of this region (29.9%) suggests that the EPS cluster was acquired via HGT. In addition to the glycosyltransferases involved in EPS synthesis, 18 other putative glycosyltransferases were also identified that were more likely involved in the synthesis of cell wall polysaccharides. All of the genes encoding glycosyltransferases in *L. gasseri* are highly conserved in *L. acidophilus* and *L. johnsonii*, with the exception of the glycosyltransferases in the EPS gene cluster, and LGAS_1543, potentially involved in biosynthesis of teichoic acid. This might provide specific surface properties to *L. gasseri* ATCC 33323.

Jacobsen et al. (53) reported that adhesion of *L. plantarum* cultures to Caco-2 epithelial cells was strain dependent. The same phenomenon was observed with strains of *L. casei* (96). In the present study, adhesion to Caco-2 by selected *L. gasseri* strains was investigated in order to assess strain diversity and to further characterize ATCC 33323. As observed for *L. plantarum* and *L. casei*, the adhesion of *L. gasseri* varied between strains (Table 4). However, two clearly distinctive groups were identified within the analyzed *L. gasseri* strains. One group, which included ATCC 33323, adhered better but, overall, the *L. gasseri* strains showed a lower ability to bind to Caco-2 monolayers compared to *L. acidophilus* NCFM.

(iv) Degradation of oxalate. A unique feature of *L. gasseri* ATCC 33323, shared only with *L. acidophilus* NCFM and *L. reuteri*, is its capability to degrade oxalate. Oxalic acid is found in dietary sources (such as coffee, tea, and chocolate); it can also be produced by the intestinal microflora from metabolic precursors, such as ascorbic acid (75). The normal western diet has an oxalate content of approximately 80 to 120 mg/day (50). A total of 10% of the oxalate consumed is normally absorbed through the intestinal tract, and ca. 50 to 80% is degraded by bacteria to CO_2 and formate, which is later metabolized or directly excreted. When oxalate-degrading bacteria are absent, intestinal oxalate absorption is increased and, consequently, urinary oxalate excretion is elevated. Oxalate at high concentrations can cause pathological disorders, including hyperoxaluria, pyridoxine deficiency, urolithiasis, renal failure, and other disorders (49).

Oxalobacter formigenes, a natural inhabitant of the GIT of vertebrates, including humans, is the best-characterized microorganism of the intestinal microbiota with an oxalate-degrading mechanism (31). In addition, the gene encoding a novel oxalyl coenzyme A decarboxylase from *Bifidobacterium lactis* DSM 10140 was identified and characterized (39). Moreover, we identified and functionally characterized an operon containing genes homologous to a formyl coenzyme A transferase gene (*frc*) and an oxalyl coenzyme A decarboxylase gene (*oxc*) in the genome of *L. acidophilus* (8). The ability of *L. gasseri* ATCC 33323 to degrade oxalate was investigated in vitro (66). The authors of that study, using reverse transcription-PCR, proved that *frc* and *oxc* genes are cotranscribed and confirmed that, as in *L. acidophilus*, *oxc* is induced by oxalate under mildly acidic (pH 5.5) conditions.

In the present study, oxalate degradation activity by *L. gasseri* ATCC 33323 was monitored and compared to selected *Lactobacillus* strains (Table 1 and Fig. 6). We observed a considerable variability in the oxalate-degrading properties of the analyzed strains. The screening allowed us to identify a number of strains degrading more than 50% of the oxalate added to the culture medium. *L. gasseri* ML3 had an oxalate-degrading activity comparable to *L. acidophilus* NCFM, while *L. gasseri* FR2, RF14, and RF81 proved to be the most active strains in oxalate degradation. No oxalate degradation was observed for *L. helveticus*, *L. jonhsonii*, and *L. gasseri* strains WD19 and ADH. *L. acidophilus* NCFM degraded 91% of total oxalate after 5 days. Although *L. gasseri* ATCC 33323 was not as effective, oxalate concentration in the supernatant decreased by 50% over the same period of time, results that are consistent with previous reports (66).

(v) Resistance to stress conditions. According to Sanderson (89), the intestinal environment results from three main factors: dietary intake, bacterial ecology, and physiology, including factors such as peristalsis and glandular secretions. A num-

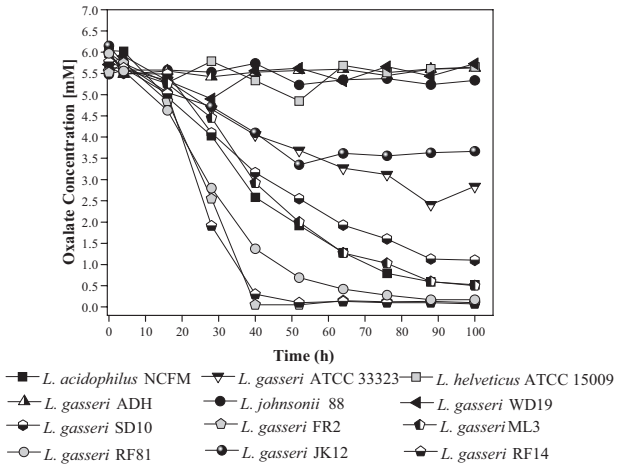


FIG. 6. Evaluation of in vitro oxalate degradation by 12 *Lactobacillus* strains.

ber of these factors restrict bacterial cell growth in the small bowel, including gastric acidity, digestive enzymes, bile salts, peristalsis, mucus, the resident commensal microflora, exfoliation of enterocytes during epithelial renewal, epithelial translocation of secretory immunoglobulin A, $CD8^+$ intraepithelial T lymphocytes, and innate host defense mechanisms mediated by gene-encoded antimicrobial peptides (reviewed by Ouellette [78]). The genome sequence of *L. gasseri* encodes the molecular chaperons and chaperonins (GroES, GroEL, DnaK, DnaJ, and GrpE; see Table S7 in the supplemental material) that protect proteins in the cytoplasm from irreversible aggregation during synthesis and stress. In addition, a number of genes contain conserved domains that can identify them as putatively involved in general stress responses (LGAS_0028, a general stress response protein CsbD; LGAS_0125 and LGAS_1247, universal stress proteins UspA), osmotic stress, heat shock, oxidative stress (including a methionine sulfoxide reductase B, LGAS_1142) and cold shock (see Table S7 in the supplemental material). Interestingly, six ORFs were found to contain the conserved domain Pfam01047 (the MarR family). The MarR family is widely distributed in nature, and the transcriptional regulators members of this family have been shown to be involved in multiple antibiotic resistance, a nonspecific resistance system including resistance to multiple antibiotics, household disinfectants, organic solvents, and oxidative stress agents (reviewed by Ellison and Miller [35]). Also of interest is the presence of a 126-amino-acid ORF containing a conserved domain involved in resistance to phenolic acids (IPR005149, LGAS_0956). *L. plantarum* encodes an inducible phenolic acid decarboxylase activity that converts these substrates into less toxic vinyl phenol derivatives (47); however, LGAS_0956 does not show significant similarity to the system encoded by *L. plantarum*. Consequently, functional studies are needed to demonstrate the involvement, if any, of this protein in the resistance to these toxic chemical compounds.

It is known that the low pH of the stomach, along with the presence of pepsin, provides an effective barrier against bacteria, inhibiting their entrance into the GIT. In the present study, we tested the effect of simulated gastric juice at pH 2.0 on the viability of 10 *Lactobacillus* strains. Wide intraspecies variability was observed between *L. gasseri* strains, with percent survival values ranging from 6.6 to 100% after 60 min of exposure and from 0 to 62.2% after 120 min (Table 4). ATCC 33323 showed moderate resistance to simulated gastric juice.

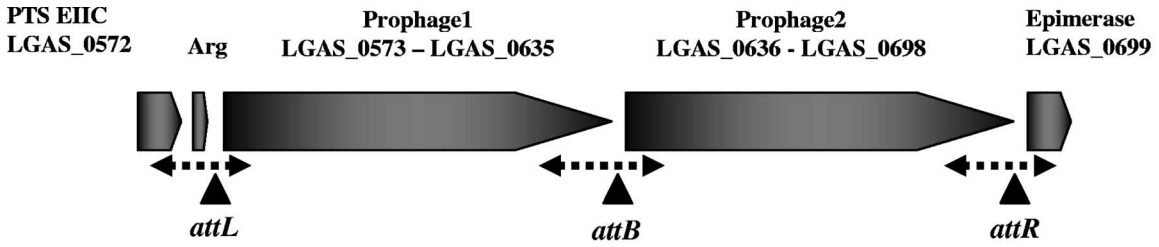
(vi) 2CRSs and other transcriptional regulators. Typically, bacterial genomes encode a sigma factor devoted to the transcriptional regulation of housekeeping genes. In addition, alternative sigma factors can control specialized regulons activated during stress, growth transitions, and morphological changes (46). *L. gasseri* encodes a housekeeping sigma factor (*rpoD*, LGAS_1126) and can choose between a pool of three alternative sigma factors (LGAS_0342; LGAS_1174, and LGAS_1483). LGAS_0342 is well conserved, and homologs are present in most *Lactobacillus* genomes. Interestingly, a BLAST search in the Microbes database at the NCBI indicated that LGAS_1174 has homologs only in a limited number of sequenced lactobacilli (*L. johnsonii* NCC 533, *L. plantarum* WCFS1, two strains of *L. delbrueckii* subsp. *bulgaricus* [ATCC BAA-365 and ATCC 11842], and *L. casei* ATCC 334) and is not present in *L. acidophilus* NCFM. More interesting, the alternative sigma factor LGAS_1483 appears to be unique to *L. gasseri* among lactic acid bacteria. Its closest BLAST hit in the Microbes database is a protein from *Bacteroides caccae* ATCC 43185 (GenBank accession number AAVM02000009; e-value 0.12; identity, 21%). *B. caccae* represents 2.8% of the total number of microbial 16S rRNA gene sequences found in a study of the colonic and fecal microbiotas of three healthy adults (33).

Gene expression levels are further modulated by the action of transcriptional regulators. Seventy putative transcriptional regulators were identified in *L. gasseri* based on the presence of conserved functional domains (see Table S8 in the supplemental material). Also observed for *L. acidophilus* (4) is that most of the identified regulators are repressors.

Only five 2CRSs were identified in the genome sequence of *L. gasseri* (Fig. 1 and see Table S9 in the supplemental material), a low number compared to other sequenced lactobacilli (*L. sakei* encodes 10 2CRSs [23], *L. acidophilus* encodes 9 2CRSs [4], *L. johnsonii* encodes 9 2CRSs [83], *L. plantarum* encodes 13 2CRSs [61], *L. casei* ATCC 334 encodes 15 2CRSs, *L. brevis* ATCC 367 encodes 9 2CRSs, and *L. delbrueckii* subsp. *bulgaricus* ATCC BAA-365 encodes 6 2CRSs [http://genome.jgi-psf.org/tre_home.html]). The function of three of *L. gasseri* 2CRSs could be inferred based on homologies to previously characterized signal transduction systems. The 2CRS composed of LGAS_0060 and LGAS_0061 appears to be part of an operon, similar to the *yycF* and *yycG* genes in *Bacillus subtilis* (37), which is essential and potentially involved in growth. ORFs LGAS_0712 - LGAS_0713 form a 2CRS similar to the bile-inducible system in *L. acidophilus* involved in resistance to bile (81). Finally, the 2CRS composed of LGAS_1410 and LGAS_1411 is similar to the 2CRS involved in acid resistance and regulation of members of the proteolytic enzyme system in *L. acidophilus* (9). In addition, we identified three orphan response regulators containing the LytTR DNA-binding motif. Other genes putatively involved in signal transduction are indicated in Table S9 in the supplemental material.

(vii) LuxS, bacteriocin, and restriction and modification (R/M) systems. Recent studies of the GIT using metagenomics have given insight into this complex microbial environment, revealing the presence of an estimated 10^{13} to 10^{14} bacterial cells present in this environment (45). Bacteria in the GIT can regulate their gene expression via cell signaling molecules in response to their surroundings. Autoinducer 2 is a signal that regulates a wide range of bacterial physiological conditions (80). While this area has been studied more extensively in pathogenic species (99), cell signaling has also recently been studied in lactobacilli such as *L. rhamnosus* GG (65), *L. reuteri* 100-23 (93), and *L. acidophilus* NCFM (19). In these studies, the *luxS* gene was inactivated, and the subsequent mutants showed lower adherence to Caco-2 cells (19) and differences in biofilm formation (65, 93) compared to the wild type. In silico analysis of the genome of *L. gasseri* ATCC 33323 revealed the *luxS* gene (LGAS_1630) and the genes encoding enzymes required for the activated methyl cycle. All of these genes shared high identity (86 to 96%) with homologues in *L. johnsonii* NCC533.

Bacteriocins are small antimicrobial peptides that are produced by gram-positive bacteria, including some lactic acid bacteria (57). Secretion of these antimicrobial peptides can kill




```
Arg tRNA AGAACAATGGTCTCCTAAACTGTAGAGGTGAGTTCGAATCTCACCGGGGTCA
attL     AGAACAATGGTCTCCTAAACTGTAGAGGTGAGTTCGAATCTCACCGGGGTCATATTGATGTAGA 600,697 bp
attB     AACGATTTGGTCTCCTAAACTGTAGAGGTGAGTTCGAATCTCACCGGGGTCATATTGATGTAGA 640,783 bp
attR     AACGATTTGGTCTCCTAAACCGTAGAGGTGAGTTCGAATCTCACCGGGGTCATAACAAAAAAGG 680,869 bp
```

FIG. 7. Schematic representation of the tandem prophages integrated in *L. gasseri* ATCC 33323. Three *att* sites (*attL*, *attB*, and *attR*), homologous to the Arg tRNA sequence sites, flanking and in-between the two tandem phages are shown.

other competing bacteria, and some are involved in cell signaling (60). *L. gasseri* ATCC 33323 does not appear to encode any putative bacteriocin peptides, as are produced by its nearest relatives *L. johnsonii* and *L. acidophilus*. In fact, no bacteriocin activity was detected in supernatants or agar cultures of *L. gasseri* ATCC 33323 (data not shown).

R/M systems function to degrade foreign DNA and are the most common systems used to degrade phage DNA. Three types of R/M systems have been described (73) and in silico analysis of *L. gasseri* ATCC 33323 reveals type I and III systems in the genome (Fig. 1). Type I R/M systems encode three subunits. Two, HsdM (LGAS_0902) and HsdS (LGAS_0903 and LGAS_0904) subunits, function for methylase activity. In addition, the HsdS subunit contains the specificity domain, with two target recognition sequences. The third subunit, HsdR (LGAS_0906) functions as the restriction unit (73). Interestingly, a phage integrase (LGAS_0904) gene was located between the two HsdS genes. The G+C content of this region (LGAS_0902 to LGAS_0906) was lower (31.6%) than that for the genome (35.26%), suggesting this could be a region where DNA was acquired by HGT. In addition, this type I R/M system in *L. gasseri* ATCC 33323 does not share any homologies with any lactobacilli of human origin, except for *L. reuteri* F275 (GenBank accession number NC_009513).

A type III R/M system is also located in the genome. Type III R/M systems are not as well characterized as type I but are composed of two subunits: Mod (LGAS_1477 and LGAS_1478) and Res (LGAS_1476). The Mod subunit is responsible for DNA recognition and methylation of the recognition site, whereas Res cleaves the DNA when bound to Mod (73). This R/M system appears to be unique to *L. gasseri* ATCC 33323, with no homologues for the complete system in any other lactic acid bacteria sequenced to date.

Analysis of prophage sequences in the *L. gasseri* genome. Genomic analyses of *L. gasseri* ATCC 33323 revealed the presence of one complete prophage sequence, LgaI. The LgaI prophage belongs to the group Sfi11-like *Siphoviridae* phage family. The prophage sequence consists of 40,086 bp, located at ~600 kb on the *L. gasseri* chromosome (Fig. 1). Interestingly, two identical copies of the prophage were integrated back-to-back on the chromosome, between bp 600641 and 640727 bp for the first copy (from LGAS_0573 to LGAS_0635) and between bp 640728 to 680814 (from LGAS_0636 to LGAS_0698) for the second (Fig. 7). Specifically, the prophage genome consists of 60 ORFs, where the first 4 ORFs and the last 56 ORFs are encoded on opposite strands. This organization is similar to the *L. gasseri* temperate bacteriophage ϕadh (3) and KC5a (GenBank accession no. DQ320509).

The prophage genome architecture is relatively typical, with a mosaic pattern of homologies with other bacteriophages, and includes the commonly found phage genome modules of lysogeny, replication, head morphogenesis, tail morphogenesis, and lysis (100). A detailed and comparative analysis of the *L. gasseri* LgaI prophage genome has been published (100), comparing the prophage sequence in various *Lactobacillus* genomes.

The phage genome sequence has seemingly integrated in a tRNA gene, specifically the Arg tRNA present at bp 600613 to 600685 on the *L. gasseri* chromosome. Three *att* sites, namely, *attL*, *attB*, and *attR*, showing homology to the Arg tRNA sequence sites are found flanking and in between the two tandem phages. This organization is consistent with a double integration of the phage, resulting in a tandem phage at a unique location. Interestingly, these attachment sites show similarity to sequences found in phigaY, another *L. gasseri* phage (105).

To confirm the integration site of each prophage and the integration of two phages in tandem, PCR primers were designed to amplify products between the prophage sequences and the flanking genes on the chromosome. Specifically, an amplicon was obtained between LGAS_0572 and LGAS_0573, including the *attL* sequence, confirming the integration of the first prophage. Also, an amplicon was obtained between LGAS_0634 and LGAS_0636, including the *attB* sequence, confirming the contiguous integration of two copies of the prophage. In addition, an amplicon was obtained between LGAS_0697 and LGAS_0699, including the *attR* sequence, confirming the integration of the second copy of the prophage, flanking LGAS_0700 (data not shown).

Restriction digest analysis of the *L. gasseri* genome by PFGE revealed the presence of a band corresponding to the region, encompassing the two prophages in tandem (Fig. 8). Specifically, SmaI sites (CCCGGG) can be found at bp 568793 to 568798 and bp 685020 to 685025 on the *L gasseri* genome sequence, thus resulting in a hypothetical SmaI digest band of 116,227 bp. This was consistent with the appearance of an ~116-kb band observed on the SmaI PFGE analysis of the *L.*

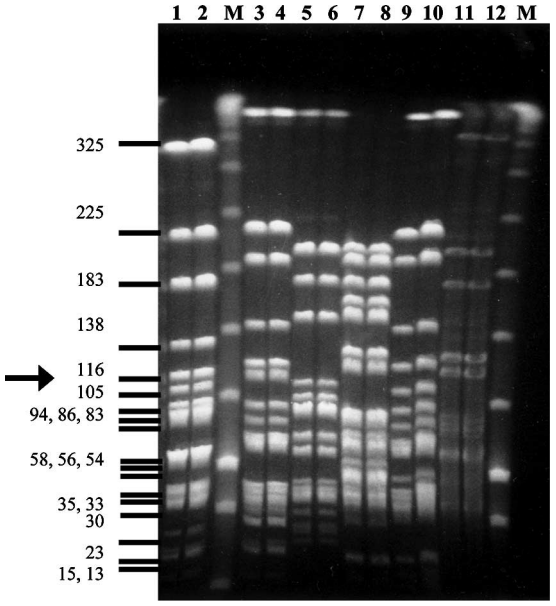


FIG. 8. PFGE patterns of *L. gasseri* strains. Lanes: 1 and 2, *L. gasseri* ATCC 33323; M, molecular weight markers; 3 and 4 *L. gasseri* ADH; 5 and 6, *L. acidophilus* NCFM; 7 and 8, *L. gasseri* JK12; 9 and 10, *L. gasseri* SD10; 11 and 12, *L. gasseri* ML3. The arrow indicates the band corresponding to 116 kb with the tandem prophage in *L. gasseri* ATCC 33323.

gasseri ATCC 33323 chromosomal DNA (Fig. 8). Although some of the other *L. gasseri* strains exhibited bands of similar size, it was not clear from this analysis whether they also carried the tandem phage organization. Nevertheless, the PFGE patterns did show diversity in the overall genome organization between the various strains of this species. Further experiments are required to investigate the functionality of these phages and the impact of the tandem organization on their life cycle.

In addition to the tandem prophages found at ~600 kb on the *L. gasseri* chromosome, a phage remnant was identified at ~1,461 kb (Fig. 1), between LGAS_1485 and LGAS_1500, with a primase (LGAS_1489), a helicase (LGAS_1490), and a major capsid protein (LGAS_1498). Interestingly, the major capsid protein shows high similarity (90% identity) to ORF lr0866, seemingly derived from a prophage present in the draft genome sequence of *L. reuteri* ATCC 55730 (GenBank accession no. ABO43797).

Conclusions. Extensive similarity at the sequence level was observed between *L. gasseri* ATCC 33323 and *L. johnsonii* NCC533 (58), with overall genome synteny and significant similarity for ca. 50% of predicted ORFs in *L. gasseri* sharing similarities to *L. johnsonii* ORFs at a level of 1e-100 and below. However, a number of unique features were identified in the genome sequence of *L. gasseri* that appear to contribute to the adaptation of the bacterium to its ecological niche, the human GIT. In addition, many of these characteristics appear to be acquired by HGT, as indicated by different G+C content regions and/or the presence of flanking transposases.

The human GIT is a complex environment that provides a variety of ecological challenges. The features of *L. gasseri* suggest that this organism is a natural part of a complex equilibrium of commensal flora in parts of the human ecosystem that participate in defense and protection of the GIT and vagina, fulfilling important functions. Despite the extensive similarity levels found at the sequence level with *L. johnsonii*, a high intraspecies variability was observed in the present study. This level of variability points out the importance of strain sequencing and in-depth studies of strain-specific genetic systems. Phenotypic traits such as carbohydrate fermentation patterns, oxalate degradation, and adhesion to intestinal epithelial cells verify the differences of probiome organisms that impact survival, association with the intestinal epithelium, immunomodulation, and interactions with the intestinal microbiota.

ACKNOWLEDGMENTS

This study was supported by the Joint Genome Institute of the Department of Energy, the North Carolina Dairy Foundation, Danisco USA, Inc., Dairy Management, Inc., and the Southeast Dairy Foods Research Center.

We thank David Mills (UC Davis) for leadership of the Lactic Acid Bacteria Genome Consortium and Paul Richardson (JGI) and Fidelity Systems for finishing and polishing the complete genome sequence for *L. gasseri* submitted to the NCBI. We thank Evelyn Durmaz for critical reading of the manuscript.

REFERENCES

1. **Altermann, E., and T. R. Klaenhammer.** 2003. GAMOLA: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes. OMICS **7:**161–169.
2. **Altermann, E., and T. R. Klaenhammer.** 2005. PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. BMC Genomics **6:**60.
3. **Altermann, E., J. R. Klein, and B. Henrich.** 1999. Primary structure and features of the genome of the *Lactobacillus gasseri* temperate bacteriophage ϕadh. Gene **236:**333–346.
4. **Altermann, E., W. M. Russell, M. A. Azcarate-Peril, R. Barrangou, et al.** 2005. Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. Proc. Natl. Acad. Sci. USA **102:**3906–3912.
5. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, et al.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
6. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
7. **Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, et al.** 2000. Gene ontology: tool for the unification of biology. Nat. Genet. **25:**25–29.
8. **Azcarate-Peril, M. A., J. M. Bruno-Barcena, H. M. Hassan, and T. R. Klaenhammer.** 2006. Transcriptional and functional analysis of oxalyl-coenzyme A (CoA) decarboxylase and formyl-CoA transferase genes from *Lactobacillus acidophilus*. Appl. Environ. Microbiol. **72:**1891–1899.
9. **Azcarate-Peril, M. A., O. McAuliffe, E. Altermann, S. Lick, et al.** 2005. Microarray analysis of a two-component regulatory system involved in acid resistance and proteolytic activity in *Lactobacillus acidophilus*. Appl. Environ. Microbiol. **71:**5794–5804.
10. **Barefoot, S. F., and T. R. Klaenhammer.** 1983. Detection and activity of lactacin B, a bacteriocin produced by *Lactobacillus acidophilus*. Appl. Environ. Microbiol. **45:**1808–1815.
11. **Bateman, A., E. Birney, L. Cerruti, R. Durbin, et al.** 2002. The Pfam protein families database. Nucleic Acids Res. **30:**276–280.
12. **Begley, M., C. G. Gahan, and C. Hill.** 2005. The interaction between bacteria and bile. FEMS Microbiol. Rev. **29:**625–651.
13. **Begley, M., C. Hill, and C. G. Gahan.** 2006. Bile salt hydrolase activity in probiotics. Appl. Environ. Microbiol. **72:**1729–1738.
14. **Begley, M., R. D. Sleator, C. G. Gahan, and C. Hill.** 2005. Contribution of three bile-associated loci, *bsh*, *pva*, and *btlB*, to gastrointestinal persistence and bile tolerance of *Listeria monocytogenes*. Infect. Immun. **73:**894–904.
15. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340:**783–795.
16. **Berger, B., R. D. Pridmore, C. Barretto, F. Delmas-Julien, et al.** 2007. Similarity and differences in the *Lactobacillus acidophilus* group identified by polyphasic analysis and comparative genomics. J. Bacteriol. **189:**1311–1321.
17. **Boekhorst, J., Q. Helmer, M. Kleerebezem, and R. J. Siezen.** 2006. Comparative analysis of proteins with a mucus-binding domain found exclusively in lactic acid bacteria. Microbiology **152:**273–280.
18. **Bron, P. A., C. Grangette, A. Mercenier, W. M. de Vos, et al.** 2004. Iden-

tification of *Lactobacillus plantarum* genes that are induced in the gastrointestinal tract of mice. J. Bacteriol. **186:**5721–5729.
19. **Buck, B. L.** 2006. Functional analysis of adhesion factors and signaling mechanisms in *Lactobacillus acidophilus* NCFM. Ph.D. thesis. North Carolina State University, Raleigh.
20. **Buck, B. L., E. Altermann, T. Svingerud, and T. R. Klaenhammer.** 2005. Functional analysis of putative adhesion factors in *Lactobacillus acidophilus* NCFM. Appl. Environ. Microbiol. **71:**8344–8351.
21. **Burkholder, P. R., and I. McVeigh.** 1942. Synthesis of vitamins by intestinal bacteria. Proc. Natl. Acad. Sci. USA **28:**285–289.
22. **Callanan, M., P. Kaleta, J. O'Callaghan, O. O'Sullivan, 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.
23. **Chaillou, S., M. C. Champomier-Verges, M. Cornet, A. M. Crutz-Le Coq, et al.** 2005. The complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K. Nat. Biotechnol. **23:**1527–1533.
24. **Charteris, W. P., P. M. Kelly, L. Morelli, and J. K. Collins.** 1998. Development and application of an in vitro methodology to determine the transit tolerance of potentially probiotic *Lactobacillus* and *Bifidobacterium* species in the upper human gastrointestinal tract. J. Appl. Microbiol. **84:**759–768.
25. **Chauviere, G., M. H. Coconnier, S. Kerneis, J. Fourniat, et al.** 1992. Adhesion of human *Lactobacillus acidophilus* strain LB to human enterocyte-like Caco-2 cells. J. Gen. Microbiol. **138**(Pt. 8)**:**1689–1696.
26. **Claesson, M. J., Y. Li, S. Leahy, C. Canchaya, et al.** 2006. Multireplicon genome architecture of *Lactobacillus salivarius*. Proc. Natl. Acad. Sci. USA **103:**6718–6723.
27. **Crowell, D. C., and T. Klaenhammer.** 1998. Microbial analysis of human intestinal flora after feeding *Lactobacillus acidophilus*. M.S. thesis. North Carolina State University, Raleigh.
28. **Dal Bello, F., and C. Hertel.** 2006. Oral cavity as natural reservoir for intestinal lactobacilli. Syst. Appl. Microbiol. **29:**69–76.
29. **De Smet, I., L. Van Hoorde, M. Vande Woestyne, H. Christiaens, et al.** 1995. Significance of bile salt hydrolytic activities of lactobacilli. J. Appl. Bacteriol. **79:**292–301.
30. **De Vuyst, L., and B. Degeest.** 1999. Heteropolysaccharides from lactic acid bacteria. FEMS Microbiol. Rev. **23:**153–177.
31. **Duncan, S. H., A. J. Richardson, P. Kaul, R. P. Holmes, et al.** 2002. *Oxalobacter formigenes* and its potential role in human health. Appl. Environ. Microbiol. **68:**3841–3847.
32. **Durbin, R., S. R. Eddy, A. Krogh, and G. Mitchison.** 1998. Markov chains and hidden Markov models in biological sequence analysis: probabilistic models of proteins and nucleic acids, p. 46–79. Cambridge University Press, Cambridge, United Kingdom.
33. **Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, et al.** 2005. Diversity of the human intestinal microbial flora. Science **308:**1635–1638.
34. **Elkins, C. A., S. A. Moser, and D. C. Savage.** 2001. Genes encoding bile salt hydrolases and conjugated bile salt transporters in *Lactobacillus johnsonii* 100-100 and other *Lactobacillus* species. Microbiology **147:**3403–3412.
35. **Ellison, D. W., and V. L. Miller.** 2006. Regulation of virulence by members of the MarR/SlyA family. Curr. Opin. Microbiol. **9:**153–159.
36. **Ermolaeva, M. D., H. G. Khalak, O. White, H. O. Smith, et al.** 2000. Prediction of transcription terminators in bacterial genomes. J. Mol. Biol. **301:**27–33.
37. **Fabret, C., and J. A. Hoch.** 1998. A two-component signal transduction system essential for growth of *Bacillus subtilis*: implications for anti-infective therapy. J. Bacteriol. **180:**6375–6383.
38. **Fanaro, S., R. Chierici, P. Guerrini, and V. Vigi.** 2003. Intestinal microflora in early infancy: composition and development. Acta Paediatr. Suppl. **91:**48–55.
39. **Federici, F., B. Vitali, R. Gotti, M. R. Pasca, et al.** 2004. Characterization and heterologous expression of the oxalyl coenzyme A decarboxylase gene from *Bifidobacterium lactis*. Appl. Environ. Microbiol. **70:**5066–5073.
40. **Fernandez, M., and M. Zuniga.** 2006. Amino acid catabolic pathways of lactic acid bacteria. Crit. Rev. Microbiol. **32:**155–183.
41. **Foster, J. W., and A. G. Moat.** 1980. Nicotinamide adenine dinucleotide biosynthesis and pyridine nucleotide cycle metabolism in microbial systems. Microbiol. Rev. **44:**83–105.
42. **Gahan, C. G., and C. Hill.** 2000. The use of listeriolysin to identify in vivo induced genes in the gram-positive intracellular pathogen *Listeria monocytogenes*. Mol. Microbiol. **36:**498–507.
43. **Ganeshkumar, N., P. M. Hannam, P. E. Kolenbrander, and B. C. McBride.** 1991. Nucleotide sequence of a gene coding for a saliva-binding protein (SsaB) from *Streptococcus sanguis* 12 and possible role of the protein in coaggregation with actinomyces. Infect. Immun. **59:**1093–1099.
44. **Gasser, F., and M. Mandel.** 1968. Deoxyribonucleic acid base composition of the genus *Lactobacillus*. J. Bacteriol. **96:**580–588.
45. **Gill, S. R., M. Pop, R. T. Deboy, P. B. Eckburg, et al.** 2006. Metagenomic analysis of the human distal gut microbiome. Science **312:**1355–1359.
46. **Gruber, T. M., and C. A. Gross.** 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu. Rev. Microbiol. **57:**441–466.
47. **Gury, J., L. Barthelmebs, N. P. Tran, C. Divies, et al.** 2004. Cloning, deletion, and characterization of PadR, the transcriptional repressor of the phenolic acid decarboxylase-encoding *padA* gene of *Lactobacillus plantarum*. Appl. Environ. Microbiol. **70:**2146–2153.
48. **Haft, D. H., J. D. Selengut, and O. White.** 2003. The TIGRFAMs database of protein families. Nucleic Acids Res. **31:**371–373.
49. **Hatch, M., and R. W. Freel.** 1995. Alterations in intestinal transport of oxalate in disease states. Scanning Microsc. **9:**1121–1126.
50. **Hoppe, B., G. von Unruh, N. Laube, A. Hesse, et al.** 2005. Oxalate degrading bacteria: new treatment option for patients with primary and secondary hyperoxaluria? Urol. Res. **33:**372–375.
51. **Hoskins, L. C., M. Agustines, W. B. McKee, E. T. Boulding, et al.** 1985. Mucin degradation in human colon ecosystems. Isolation and properties of fecal strains that degrade ABH blood group antigens and oligosaccharides from mucin glycoproteins. J. Clin. Investig. **75:**944–953.
52. **Hugenholtz, J., W. Sybesma, M. N. Groot, W. Wisselink, et al.** 2002. Metabolic engineering of lactic acid bacteria for the production of nutraceuticals. Antonie van Leeuwenhoek **82:**217–235.
53. **Jacobsen, C. N., V. Rosenfeldt-Nielsen, A. E. Hayford, P. L. Moller, et al.** 1999. Screening of probiotic activities of forty-seven strains of *Lactobacillus* spp. by in vitro techniques and evaluation of the colonization ability of five selected strains in humans. Appl. Environ. Microbiol. **65:**4949–4956.
54. **Johnson, J. L., C. S. Phelps, C. S. Cummins, J. London, et al.** 1980. Taxonomy of the *Lactobacillus acidophilus* group. Int. J. Syst. Bacteriol. **30:**53–63.
55. **Kilic, A. O., L. Tao, Y. Zhang, Y. Lei, et al.** 2004. Involvement of *Streptococcus gordonii* beta-glucoside metabolism systems in adhesion, biofilm formation, and in vivo gene expression. J. Bacteriol. **186:**4246–4253.
56. **Klaenhammer, T., E. Altermann, F. Arigoni, A. Bolotin, et al.** 2002. Discovering lactic acid bacteria by genomics. Antonie van Leeuwenhoek **82:**29–58.
57. **Klaenhammer, T. R.** 1988. Bacteriocins of lactic acid bacteria. Biochimie **70:**337–349.
58. **Klaenhammer, T. R., R. Barrangou, B. L. Buck, M. A. Azcarate-Peril, et al.** 2005. Genomic features of lactic acid bacteria effecting bioprocessing and health. FEMS Microbiol. Rev. **29:**393–409.
59. **Kleeman, E. G., and T. R. Klaenhammer.** 1982. Adherence of *Lactobacillus* species to human fetal intestinal cells. J. Dairy Sci. **65:**2063–2069.
60. **Kleerebezem, M.** 2004. Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis. Peptides **25:**1405–1414.
61. **Kleerebezem, M., J. Boekhorst, R. van Kranenburg, D. Molenaar, et al.** 2003. Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proc. Natl. Acad. Sci. USA **100:**1990–1995.
62. **Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. **305:**567–580.
63. **Kullen, M. J., R. B. Sanozky-Dawes, D. C. Crowell, and T. R. Klaenhammer.** 2000. Use of the DNA sequence of variable regions of the 16S rRNA gene for rapid and accurate identification of bacteria in the *Lactobacillus acidophilus* complex. J. Appl. Microbiol. **89:**511–516.
64. **Lauer, E., and O. Kandler.** 1980. *Lactobacillus gasseri* sp. nov., a new species of the subgenus *Thermobacterium*. Zentralbl. Bakteriol. Mikrobiol. Hyg. Abt. 1 Orig. C **1:**75–78.
65. **Lebeer, S., S. C. De Keersmaecker, T. L. Verhoeven, A. A. Fadda, et al.** 2007. Functional analysis of *luxS* in the probiotic strain *Lactobacillus rhamnosus* GG reveals a central metabolic role important for growth and biofilm formation. J. Bacteriol. **189:**860–871.
66. **Lewanika, T. R., S. J. Reid, V. R. Abratt, G. T. Macfarlane, et al.** 2007. *Lactobacillus gasseri* Gasser AM63^T degrades oxalate in a multistage continuous culture simulator of the human colonic microbiota. FEMS Microbiol. Ecol. **61:**110–120.
67. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
68. **Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, et al.** 2006. Comparative genomics of the lactic acid bacteria. Proc. Natl. Acad. Sci. USA **103:**15611–15616.
69. **Martin, R., S. Langa, C. Reviriego, E. Jiminez, et al.** 2003. Human milk is a source of lactic acid bacteria for the infant gut. J. Pediatr. **143:**754–758.
70. **McAuliffe, O., R. J. Cano, and T. R. Klaenhammer.** 2005. Genetic analysis of two bile salt hydrolase activities in *Lactobacillus acidophilus* NCFM. Appl. Environ. Microbiol. **71:**4925–4929.
71. **Mulder, N. J., R. Apweiler, T. K. Attwood, A. Bairoch, et al.** 2007. New developments in the InterPro database. Nucleic Acids Res. **35:**D224–D228.
72. **Munson, M. A., A. Banerjee, T. F. Watson, and W. G. Wade.** 2004. Molecular analysis of the microflora associated with dental caries. J. Clin. Microbiol. **42:**3023–3029.
73. **Murray, N. E.** 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol. Mol. Biol. Rev. **64:**412–434.

74. **Murty, V. L., F. J. Downs, and W. Pigman.** 1978. Rat-colonic, mucus glycoprotein. Carbohydr. Res. **61:**139–145.
75. **Ogawa, Y., T. Miyazato, and T. Hatano.** 2000. Oxalate and urinary stones. World, J. Surg. **24:**1154–1159.
76. **Oh, K. W., M. J. Kim, H. Y. Kim, B. Y. Kim, et al.** 2005. Enzymatic characterization of a maltogenic amylase from *Lactobacillus gasseri* ATCC 33323 expressed in *Escherichia coli*. FEMS Microbiol. Lett. **252:**175–181.
77. **O'Hara, A. M., and F. Shanahan.** 2006. The gut flora as a forgotten organ. EMBO Rep. **7:**688–693.
78. **Ouellette, A. J.** 2004. Defensin-mediated innate immunity in the small intestine. Best Pract. Res. Clin. Gastroenterol. **18:**405–419.
79. **Pedersen, A. G., L. J. Jensen, S. Brunak, H. H. Staerfeldt, et al.** 2000. A DNA structural atlas for *Escherichia coli*. J. Mol. Biol. **299:**907–930.
80. **Pei, D., and J. Zhu.** 2004. Mechanism of action of S-ribosylhomocysteinase (LuxS). Curr. Opin. Chem. Biol. **8:**492–497.
81. **Pfeiler, E. A., M. A. Azcarate-Peril, and T. R. Klaenhammer.** 2007. Characterization of a novel bile-inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus*. J. Bacteriol. **189:**4624–4634.
82. **Piddock, L. J.** 2006. Multidrug-resistance efflux pumps: not just for resistance. Nat. Rev. Microbiol. **4:**629–636.
83. **Pridmore, R. D., B. Berger, F. Desiere, D. Vilanova, et al.** 2004. The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. Proc. Natl. Acad. Sci. USA **101:**2512–2517.
84. **Reeds, P. J., and D. G. Burrin.** 2001. Glutamine and the bowel. J. Nutr. **131:**2505S–2508S, 2523S–2524S.
85. **Reuter, G.** 2001. The *Lactobacillus* and *Bifidobacterium* microflora of the human intestine: composition and succession. Curr. Issues Intest. Microbiol. **2:**43–53.
86. **Roos, S., and H. Jonsson.** 2002. A high-molecular-mass cell surface protein from *Lactobacillus reuteri* 1063 adheres to mucus components. Microbiology **148:**433–442.
87. **Ruas-Madiedo, P., M. Gueimonde, A. Margolles, C. G. de los Reyes-Gavilan, et al.** 2006. Exopolysaccharides produced by probiotic strains modify the adhesion of probiotics and enteropathogens to human intestinal mucus. J. Food Prot. **69:**2011–2015.
88. **Sampson, J. S., S. P. O'Connor, A. R. Stinson, J. A. Tharpe, et al.** 1994. Cloning and nucleotide sequence analysis of *psaA*, the *Streptococcus pneumoniae* gene encoding a 37-kilodalton protein homologous to previously reported *Streptococcus* sp. adhesins. Infect. Immun. **62:**319–324.
89. **Sanderson, I. R.** 1999. The physicochemical environment of the neonatal intestine. Am. J. Clin. Nutr. **69:**1028S–1034S.
90. **Sleator, R. D., H. H. Wemekamp-Kamphuis, C. G. Gahan, T. Abee, et al.** 2005. A PrfA-regulated bile exclusion system (BilE) is a novel virulence factor in *Listeria monocytogenes*. Mol. Microbiol. **55:**1183–1195.
91. **Smith, S. I., A. J. Aweh, A. O. Coker, K. O. Savage, et al.** 2001. Lactobacilli in human dental caries and saliva. Microbios **105:**77–85.
92. **Stalhammar-Carlemalm, M., T. Areschoug, C. Larsson, and G. Lindahl.** 1999. The R28 protein of *Streptococcus pyogenes* is related to several group B streptococcal surface proteins, confers protective immunity and promotes binding to human epithelial cells. Mol. Microbiol. **33:**208–219.
93. **Tannock, G. W., S. Ghazally, J. Walter, D. Loach, et al.** 2005. Ecological behavior of *Lactobacillus reuteri* 100-23 is affected by mutation of the *luxS* gene. Appl. Environ. Microbiol. **71:**8419–8425.
94. **Tanskanen, E. I., D. L. Tulloch, A. J. Hillier, and B. E. Davidson.** 1990. Pulsed-field gel electrophoresis of SmaI digests of lactococcal genomic DNA, a novel method of strain identification. Appl. Environ. Microbiol. **56:**3105–3111.
95. **Tatusov, R. L., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, et al.** 2003. The COG database: an updated version includes eukaryotes. BMC Bioinform. **4:**41.
96. **Tuomola, E. M., and S. J. Salminen.** 1998. Adhesion of some probiotic and dairy *Lactobacillus* strains to Caco-2 cell cultures. Int. J. Food Microbiol. **41:**45–51.
97. **van de Guchte, M., S. Penaud, C. Grimaldi, V. Barbe, et al.** 2006. The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. Proc. Natl. Acad. Sci. USA **103:**9274–9279.
98. **van Pijkeren, J. P., C. Canchaya, K. A. Ryan, Y. Li, et al.** 2006. Comparative and functional analysis of sortase-dependent proteins in the predicted secretome of *Lactobacillus salivarius* UCC118. Appl. Environ. Microbiol. **72:**4143–4153.
99. **Vendeville, A., K. Winzer, K. Heurlier, C. M. Tang, et al.** 2005. Making "sense" of metabolism: autoinducer-2, LuxS, and pathogenic bacteria. Nat. Rev. Microbiol. **3:**383–396.
100. **Ventura, M., C. Canchaya, V. Bernini, E. Altermann, et al.** 2006. Comparative genomics and transcriptional analysis of prophages identified in the genomes of *Lactobacillus gasseri*, *Lactobacillus salivarius*, and *Lactobacillus casei*. Appl. Environ. Microbiol. **72:**3130–3146.
101. **Walker, A., A. Cerdeno-Tarraga, and S. Bentley.** 2006. Faecal matters. Nat. Rev. Microbiol. **4:**572–573.
102. **Wall, R., G. Fitzgerald, S. Hussey, T. Ryan, et al.** 2007. Genomic diversity of cultivable *Lactobacillus* populations residing in the neonatal and adult gastrointestinal tract. FEMS Microbiol. Ecol. **59:**127–137.
103. **Ward, R. E., M. Ninonuevo, D. A. Mills, C. B. Lebrilla, et al.** 2006. In vitro fermentation of breast milk oligosaccharides by *Bifidobacterium infantis* and *Lactobacillus gasseri*. Appl. Environ. Microbiol. **72:**4497–4499.
104. **Xu, J., and J. I. Gordon.** 2003. Inaugural article: honor thy symbionts. Proc. Natl. Acad. Sci. USA **100:**10452–10459.
105. **Yokoi, K. J., M. Shinohara, N. Kawahigashi, K. Nakagawa, et al.** 2005. Molecular properties of the two-component cell lysis system encoded by prophage phigaY of *Lactobacillus gasseri* JCM 1131T: cloning, sequencing, and expression in *Escherichia coli*. Int. J. Food Microbiol. **99:**297–308.