[Frontiers in Bioscience 14, 1362-1386, January 1, 2009]

Genomic features of *Lactobacillus* species

Yong Jun Goh, Todd R. Klaenhammer

Department of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, NC 27695

TABLE OF CONTENTS

1. ABSTRACT

As member of the lactic acid bacteria (LAB), the genus *Lactobacillus* represents a diverse number of species that play significant roles in the biopreservation of foods and commensals common within the human gastrointestinal (GI) tract. Certain species of *Lactobacillus*, particularly those of human origin, have been used as probiotic bacteria due to their health-promoting effects. A recent explosion of genomic information on lactobacilli has expanded our knowledge of metabolic capabilities and key gene features that are predicted to play important roles in niche adaptation and function. This review provides an overview of probiotic-related genome features and functional genomic studies that have linked genes to traits. Interspecies heterogeneity and niche-specialized adaptation among lactobacilli, as revealed by comparative genome analysis, are also discussed.

2. INTRODUCTION

The genus *Lactobacillus* belongs to the phylum *Firmicutes*, class *Bacilli*, order *Lactobacillales*, and family *Lactobacillaceae*, with the genus *Pediococcus* as its closest relative within the same family. *Lactobacillus* species are phylogenetically diverse, with over 100 species documented to date (1). *Lactobacillus* thrive mostly in carbohydrate-rich environments, including milk, meat, cereals, plants, and mucosal surfaces (oral, gastrointestinal [GI] tract and reproductive tracts) of animals and humans, reflecting their fastidious nutritional requirements. They are typically low G + C Gram-positive, non-spore-forming rods or coccobacilli, catalase negative, anaerobic or aerotolerant, and acid tolerant. Lactobacilli are members of lactic acid bacteria (LAB), a broadly defined family of microorganisms that ferment various hexoses into primarily lactic acid. They are chemo-organotrophs, strictly fermentative, and are biochemically subdivided to either

Table 1. *Lactobacillus* genome sequencing progress to date (January 2008)

***Lactobacillus* species**	**PG[1]**	**Strain**	**[Origin] Application**	**Genome size (Mb)**	**Institution**	**Genbank accession no./Status**	**Ref.**
acidophilus	aci	NCFM	[Human] probiotic	1.99	NC State Univ., California Poly. State Univ., USA	CP000033	20
brevis	bre	ATCC 367	Beer, sourdough	2.34	JGI., LABGC, Fidelity Systems Inc., USA	CP000416	10
casei	cas	ATCC 334	[Swiss cheese] dairy	2.92	JGI., LABGC, Fidelity Systems Inc., USA	CP000423	10
		Shirota	Probiotic	3.04	Yakult, Japan	Not public	115
		DN-114 001	Dairy, probiotic	3.14	Danone Vitapole, INRA, France	Not public	115
		BL23	Probiotic	~2.6	INRA/CNRS, Caen Univ., France	In progress	115
delbrueckii subsp. *bulgaricus*	aci	ATCC 11842	[Bulgarian yogurt] yogurt	1.86	INRA, Genoscope, France	CR954253	11
		ATCC BAA-365	Yogurt, cheese	1.86	JGI., LABGC, Fidelity Systems Inc., USA	CP000412	10
		DN-100 107	Dairy, probiotic	2.13	Danone Vitapole, France	Not public	115
fermentum	reu	IFO3956	Plant	2.1	Azabu University, Japan	Completed	
gasseri	aci	ATCC 33323	[Human] probiotic	1.89	JGI., LABGC, Fidelity Systems Inc., USA	CP000413	10
helveticus	aci	DPC4571	[Swiss cheese] cheese, adjunct	2.08	Teagasc, Moorepark Food Research Center, Ireland	CP000517	12
		CNRZ32	Cheese flavor	~2.6	Univ. of Wisconsin, USA	In progress	115
		CM4	Functional foods	2.03	Calpis, Kitasato Univ., Japan	Not public	115
johnsonii	aci	NCC533	[Human] probiotic	1.99	Nestle Research Center, Switzerland	AE017198	19
plantarum	plan	WCFS1	[Human]	3.31	Wageningen Centre for Food Sciences, The Netherlands	AL935263	9
reuteri	reu	F275	[Human]	2.0	Univ. of Otaga, New Zealand; JGI, USA	CP000705	NCBI
		100-23	[Rodent gut]	2.17	Univ. of Otaga, New Zealand; JGI, USA	In progress	115
		DSM20016	[Human]	NA[2]	Univ. of Otaga, New Zealand; JGI, USA	In progress	115
		JCM1112	[Human]	2.04	Azabu University, Japan	Completed	
		ATCC 55730	[Human] probiotic	~2.0	BioGaia, Swedish Univ. of Agricultural Science, Sweden	Not public	115
rhamnosus	cas	HN001	Cheese	~2.4	New Zealand Dairy Board, New Zealand	In progress	115
sakei subsp. sakei	sak	23K	[French sausage] meat	1.88	INRA, France	CR936503	17
salivarius subsp. *salivarius*	sal	UCC118	[Human] probiotic	1.83	Univ. College Cork, Ireland	CP000233	13

[1] Phylogenetic group: aci, *L. acidophilus* group; bre, *L. brevis* group; cas, *L. casei* group; plan, *L. plantarum* group; reu, *L. reuteri* group; sak, *L. sakei* group, sal, *L. salivarius* group (1) [2] NA, not available

homofermentative (produce mainly lactic acid via the glycolytic Embden-Meyerhof pathway [EMP]), heterofermentative (produce lactic acid, acetic acid, formic acid, carbon dioxide and ethanol, via the phosphoketolase pathway [PKP]), or facultative heterofermentative (ferment hexoses via the EMP and pentoses via the PKP) (2).

Historically as one of the most important groups of microorganisms used in the biopreservation of foods, many *Lactobacillus* species have long been associated with the production of fermented foods including dairy products, vegetables, meat, and sourdough for which their desirable rapid acidification also contributes to flavor, texture, and nutrition (3). More recently, some *Lactobacillus* species are recognized as transient (allochthonous) and natural (autochthonous) gut inhabitants that contribute to intestinal health. Despite their low dominance in the GI microecology, they represent an important component of the microbiota residing in the small intestine. Consequently, certain strains of *Lactobacillus*, particularly those of human origin, have been exploited as probiotics—live microorganisms which, when administered in adequate amounts confer a health benefit on the host (4). The concept of consuming viable microorganisms for the improvement of health was initially introduced in 1908, when Élie Metchnikoff (1845-1916), a Russian immunologist and a Nobel Laureate, proposed that the consumption of fermented milk containing *Lactobacillus* promotes health and longevity by displacement of putrefactive intestinal organisms. In recent decades, accumulating evidence has supported the health-promoting roles of probiotic bacteria, including maintenance of intestinal integrity (5), pathogen exclusion (6), and immunomodulation of host (7, 8). In today's society, factors such as the emergence of multidrug-resistant pathogens, GI disorders, and increasing consumer demands

for natural remedies in health maintenance, have sparked considerable interest in probiotic cultures.

The year 2003 marked the first public release of a *Lactobacillus* complete genome from *Lactobacillus plantarum* (9), followed by a gradual increase in the number of lactobacilli genomes reported in subsequent years. Notably, the collaboration between the Lactic Acid Bacteria Genome Consortium (LABGC) and the Joint Genome Institute (JGI) from the US Department of Energy has released eleven LAB genomes simultaneously, four of which belong to *Lactobacillus* species (10). To date (January 2008), a total of fourteen *Lactobacillus* genomes are completed, with several more underway (Table 1), covering species from different phylogenetic groups (Figure 1) that are associated with various ecological niches and industrial functionalities. This review provides the current state of post-genome analysis on *Lactobacillus* genomes, emphasizing key genome features and related functional studies that are aimed at elucidating the mechanisms of probiotic actions of commensal and probiotic lactobacilli.

3. GENOME OVERVIEW

3.1. Genome statistics

The primary features of twelve completed *Lactobacillus* genomes are summarized in Table 2. The majority of the genomes ranges from 1.8 to 2 Mb, except for *L. casei* and *L. plantarum*, harboring larger chromosomes of 2.9 Mb and 3.3 Mb, respectively, similar in size to *Enterococcus* and *Listeria* species. Plasmids constitute about 14% of the total genome of *L. salivarius* with the presence of a 242-kb megaplasmid, pMP118, and two smaller plasmids. The overall GC content ranges from 33% to 37% for most lactobacilli to an unexpected 49.7% for *L. delbrueckii* subsp. *bulgaricus*. Interestingly, the unusually high GC content in *L. bulgaricus* results from an overall high %GC at the third codon position (65%), as opposed to the corresponding low %GC at the third codon in the phylogenetically related *L. acidophilus* (25%) and *L. johnsonii* (24.4%) (11). In addition, *L. bulgaricus* has the highest number of rRNA operons and tRNA genes among all sequenced prokaryotes of similar genome size (11). Although the dairy processing environment is a major reservoir for bacteriophages, the dairy-associated *L. helveticus* and *L. bulgaricus* are rarely attacked by phages and are the only sequenced strains that do not harbor prophage or phage remnants (11, 12). Nevertheless, both species have a large number of pseudogenes that account for 11% to 17% of their protein-coding genes, compared to only ~1% in *L. plantarum*, even though the latter encodes almost twice the proteome size. Meanwhile, in *L. salivarius*, nearly one-third of its pseudogenes are located on pMP118 (13). The genome of *L. helveticus* is saturated with over 200 copies of IS elements, representing among the highest number and greatest diversity among all microbial genomes sequenced to date (12).

3.2. Metabolic and biosynthetic capabilities

Carbohydrate fermentation represents the principal source of energy production in lactobacilli via the glycolytic EMP pathway or/and the PKP. Transcriptional studies and *in silico* prediction of highly expressed genes based on codon adaptation index further indicated that genes for glycolytic enzymes were among the most highly expressed in homofermentative and facultative heterofermentative lactobacilli (9, 14). *In silico* genome analysis confirmed that all species except the obligate heterofermentative *L. brevis* and *L. reuteri* F275 (*L. reuteri*) have a complete EMP pathway. The presence of an intact PKP pathway in *L. brevis* genome is in agreement with its classification. Interestingly, our preliminary examination of the *L. reuteri* genome data found only a partial PKP pathway, whereas the genome of *L. reuteri* ATCC 55730 encodes both functional EMP and PKP pathways (15). *L. salivarius*, which historically was regarded as homofermentative, possesses a functional PKP complemented by genes encoded on pMP118, and its facultative heterofermentative metabolism was later confirmed (13, 16). In addition, *L. salivarius* has all key enzymes for gluconeogenesis (also partly pMP118-encoded) that may be activated during glucose starvation. Thus, the ability to assimilate ribose, which may be a common carbon source in the GI tract due to RNA degradation, combined with gluconeogenesis from pyruvate, reflects an adaptation of *L. salivarius* to pentose utilization and is likely to confer competitive advantage in the GI environment (13). Meanwhile, the genome of the meat-fermenting *L. sakei* is equipped with the ability to obtain energy via salvage pathways of purine nucleosides inosine and adenosine, and amino acid catabolism via the arginine deiminase (ADI) pathway (17). *L. sakei* also has key enzymes for glycolytic methyglyoxal bypass that may be crucial in response to glucose starvation and regulation of carbon flux. These alternative energy production pathways are uncommon in LAB, and are apparently specialized for the survival and colonization of *L. sakei* on raw meat environments where carbohydrates are limiting (17).

One major characteristic of *Lactobacillus* species is their complex nutritional requirements and predominance of transporters in the genome. Accordingly, genome analyses confirmed that the majority of the species are highly auxotrophic, with some at intermediate auxotrophy, such as *L. plantarum* and *L. salivarius*. All species lack key enzymes for biosynthesis of most if not all vitamins and cofactors, which explains the requirement of these components in minimal growth medium. It has been reported that *L. reuteri* CRL1098 is capable of synthesizing cobalamin (vitamin B12), which is a unique feature among lactobacilli (18). Concomitantly, only the *L. reuteri* genome was found to encode putative enzymes for cobalamin synthesis, of which the closest orthologs were identified in *Listeria*, *Yersinia*, *Salmonella*, and *Streptococcus sanguinis*.

With the exception of *L. plantarum*, which has complete biosynthetic pathways for all amino acids except branched-chain amino acids (9), most *Lactobacillus* species lack *de novo* synthetic capability for various amino acids. Meanwhile, in some species, some amino acids could be generated via predicted interconvertion from existing

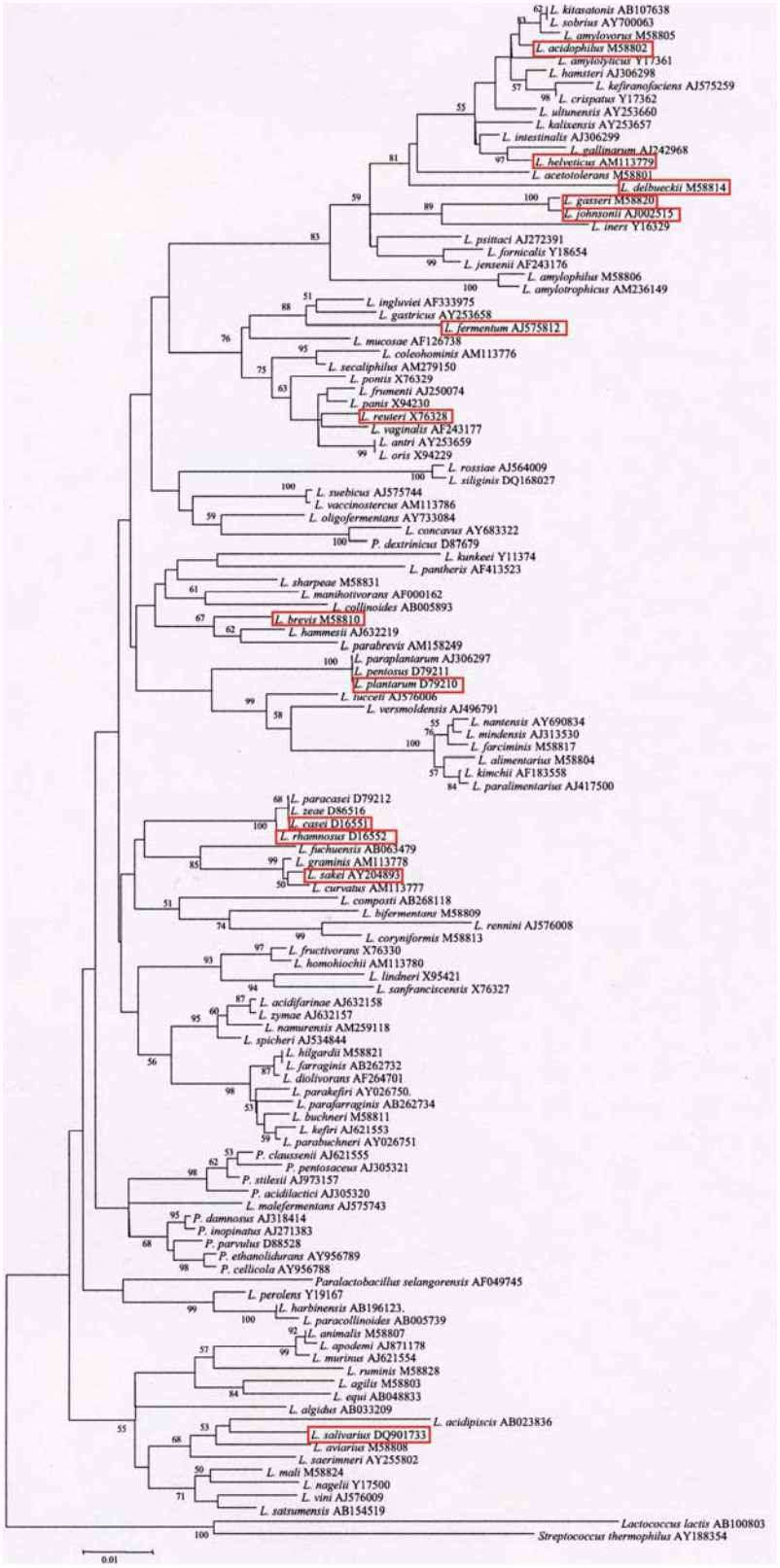

Figure 1. Phylogenetic tree of the Family *Lactobacillaceae* based on 16S rRNA gene sequence. Species by which genomes have been sequenced to date (January 2008) were outlined in red. Adapted with permission from (1).

Table 2. General features of *Lactobacillus* genomes

Species/strains	Genome size in bp [with plasmid(s)]	%GC	No. of proteins[1]	Plasmid [no. of proteins]	rRNA operons / tRNA	Pseudo-gene	Intact Prophage [phage remnant]	IS element [family]
Human isolates/probiotic strains								
L. acidophilus NCFM	1,993,560	34.7	1,862	None	4 / 61	NA	[3]	17 [7]
L. gasseri ATCC 33323	1,894,360	35.3	1,810	None	6 / 75	48	1 [1]	13 [?]
L. johnsonii NCC533	1,992,676	34.6	1,821	None	6 / 79	NA	2 [1]	14 [3]
L. plantarum WCFS1	3,308,274 [3,348,625]	44.4	3,009	1.9-kb pWCFS101 [3] 2.4-kb pWCFS102 [4] 36-kb pWCFS103 [43]	5 / 62	39	2 [2]	15 [2]
L. reuteri F275	1,999,618	38.9	1,900	0	6 / 68	39	2?	
L. salivarius subsp. *salivarius* UCC118	1,827,111 [2,133,977]	33	1,717	20-kb pSF118-20 [27] 44-kb pSF118-44 [51] 242-kb pMP118 [222]	7 / 78	73	2 [2]	37 [8]
Dairy/food fermentation strains								
L. brevis ATCC 367	2,291,220 [2,340,228]	46.1	2,185	13-kb Plasmid 1 [11] 36 kb Plasmid 2 [22]	5 / 65	50	1	
L. casei ATCC 334	2,895,264 [2,924,325]	46.6	2,751	29-kb Plasmid 1 [20]	5 / 59	82	1	
L. delbrueckii subsp. *bulgaricus* ATCC 11842	1,864,998	49.7	1,562	None	9 / 95	270	0	56 [?]
L. delbrueckii subsp. *bulgaricus* ATCC BAA-365	1,856,951	49.7	1,721	None	9 / 98	192	0	
L. helveticus DPC4571	2,080,931	37.7	1,610	None	4 / 61	217	0	213 [21]
L. sakei subsp. *sakei* 23K	1,884,661	41.3	1,879	None	7 / 63	30	[1]	12 [4]

[1] chromosomally encoded only

pathways (19, 20). The biosynthetic deficiencies are apparently compensated by the presence of a wide array of amino acid/peptide uptake systems and peptidases encoded in most of the genomes to acquire and utilize exogenous nitrogen sources. In numerous cases, the genes encoding the transporters, peptidases and genes for peptide metabolism were found organized in operons (19, 20). Both probiotic lactobacilli *L. acidophilus* and *L. johnsonii* possess putative cell wall-bound proteinases that are likely to participate in degradation of polypeptides from food substrates in the host diet. It has also been proposed that these proteinases may be involved in proteolytic cleavage of mucin glycoproteins (19). The *L. johnsonii* genome encodes a pair of tandem aminopeptidases (LJ0176/LJ0178) with exclusive homology to corresponding enzymes in other GI lactobacilli, streptococci, *Bifidobacterium longum* and *Bacteroides fragilis*, indicating selective pressure from peptide structures commonly encountered within the GI environment (19). Analysis of the *L. plantarum* genome revealed a 25-kb nonribosomal peptide synthesis (NRPS) gene cluster that is unique among the LAB. A similar biosynthetic machinery has been shown to produce peptide-like compounds with diverse structures and compositions, some of which displaying pharmaceutical values (9).

Presumably as a result of adaptation to the protein-rich milk environment, *L. bulgaricus* and *L. helveticus* have limited functional amino acid biosynthetic pathways (11, 12). *L. bulgaricus* has a proteinase gene that is important for casein degradation during growth in milk environment (11). It has been known that a mutually beneficial relationship exists between *L. bulgaricus* and *Streptococcus thermophilus* that stimulate the growth and acidification of both species during milk fermentation, a phenomenon known as protocooperation. The importance of this proteinase in protocooperation with *S. thermophilus* has been suggested, but remains to be determined, as the latter does not possess a proteinase, but is capable of synthesizing all amino acids (21). Meanwhile, novel peptidases were identified in *L. helveticus* that have a predicted specificity for proline-rich peptides derived from casein, as well as endopeptidases that may play a crucial role in reducing bitter defects during cheese ripening (12).

In general, the distribution of genes involved in amino acid or carbohydrate transport and metabolism varies among *Lactobacillus* species (Figure 2). The majority of probiotic species have a larger proportion of genes devoted to carbohydrate catabolism. In terms of carbohydrate acquisition, phosphoenolpyruvate (PEP)-dependent phosphotransferase systems (PTS) are

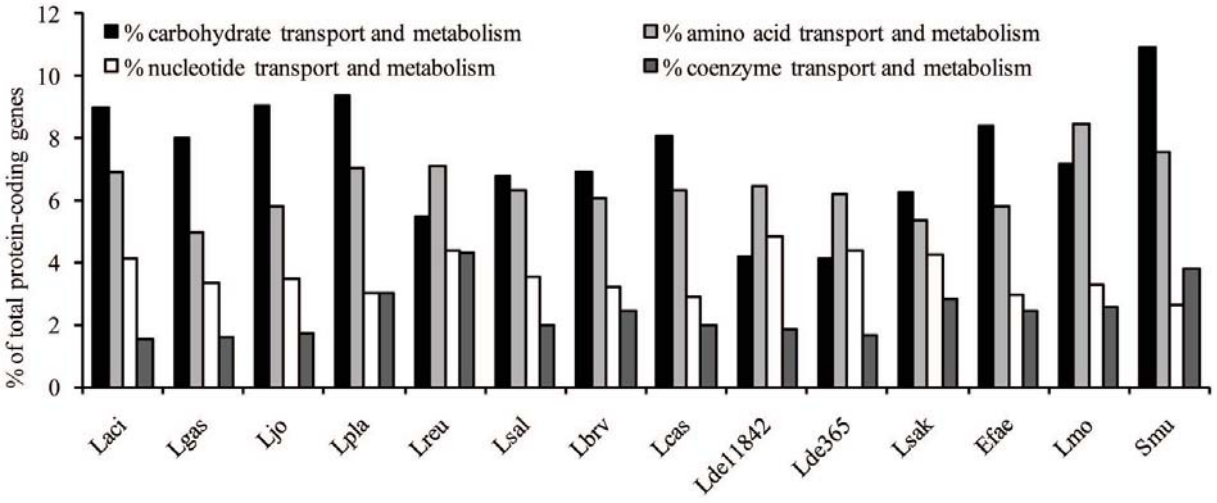

Figure 2. Distribution of putative genes involved in major metabolic and biosynthetic functions. Genes were classified based on cluster of orthologous groups (COG) functions, obtained from IMG v2.4 released by JGI (http://img.jgi.doe.gov/pub/doc/using_index.html). Laci, *L. acidophilus* NCFM; Lgas, *L. gasseri* ATCC 33323; Ljo, *L. johnsonii* NCC533; Lpla, *L. plantarum* WCFS1; Lreu, *L. reuteri* F275; Lsal, *L. salivarius* UCC118; Lbrv, *L. brevis* ATCC 367; Lcas, *L. casei* ATCC 334; Lde11842, *L. bulgaricus* ATCC 11842; Lde365, *L. bulgaricus* ATCC BAA-365; Lsak, *L. sakei* 23K; Efae, *Enterococcus faecalis* V583; Lmo, *Listeria monocytogenes* EGD-e, and Smu, *Streptococcus mutans* UA159.

predominant in these GI-associated lactobacilli, except *L. reuteri* (see below) (9, 13, 19, 20), and the dairy strain *L.* casei ATCC 334. The genome of each of these species encodes an average of 20 to 30 PTS transporters, plus several ATP-dependent binding cassette (ABC) sugar transporters and permeases of the major facilitator superfamily (MFS). Collectively, these sugar uptake systems will provide specificities for a wide range of mono- and disaccharides, and to a lesser extent, trisaccharides. Such a rich and redundant repertoire of sugar transporters implicates their ability to scavenge simple carbohydrates that are relatively abundant in the upper GI tract. Of particular interest is *L. plantarum*, which detailed genome analyses further confirm its versatility for adaptation with elaborate sets of genes that allow utilization of various niche-specific sugar substrates. Notably, some of the genes involved in sugar uptake, metabolism, and regulatory functions were encoded in a 213-kb "lifestyle adaptation island" (9). Recently, nine gene clusters were identified in the *L. plantarum* genome that encode exclusively cell surface proteins (22). These gene products presumably form cell surface protein complexes that are involved in the acquisition and/or utilization of plant-derived oligo- or polysaccharides. The orthologs of these gene clusters are limited to several LAB species and various plant-associated Gram-positive microorganisms, but are absent in the GI-associated *L. acidophilus* group and streptococci, further implying the competitive fitness of *L. plantarum* as free-living microorganisms in diverse environments beyond the GI tract.

L. bulgaricus, L. helveticus, and *L. sakei*, on the other hand, possess much fewer sugar transporters, possibly due to limited carbohydrate sources (i.e., meat) or diversity (i.e., milk) in their habitats. In *L. bulgaricus*, genes coding for the uptake and hydrolysis of the preferred milk sugar, lactose, namely lactose permease and beta-galactosidase are present (11). However, due to the lack of key enzymes for galactose utilization, only the glucose moiety of the internalized lactose was metabolized, while most of the galactose is excreted to extracellular medium by an antiporter mechanism (23). It has established previously that PTSs are rare in heterofermenters such as *L. reuteri* and *L. brevis*, with the exception of strains that possess a complete glycolytic pathway (24-26). This is because the lower yield of PEP gained from PKP is generally not sufficient for both sugar transport and biosynthetic processes (27). Thus, sugar uptake in these microorganisms is driven primarily by secondary transport systems. Consistently, no intact PTS (enzyme IIABC) was identified in the *L. reuteri* and *L. brevis* genomes; instead, both organisms encode numerous putative secondary sugar transporters with some belonging to the MFS.

4. GENOME FEATURE HIGHLIGHTS AND FUNCTIONAL GENOMICS

4.1. Environmental response and adaptation

4.1.1. Transcriptional regulation

L. plantarum encodes a large number of putative regulatory genes (262 genes or 8.5% of total proteins) which is comparable to those observed only in pathogenic microorganisms *Pseudomonas aeruginosa* (8.4%) and *Listeria monocytogenes* (7.3%) (9). Among these genes, three encode for a primary sigma factor RpoD and two alternative sigma factors, RpoN and SigH, respectively. Thirteen two-component regulatory systems (2CRS) were identified in the *L. plantarum* genome (9), only second to the fifteen 2CRS present in the *L. casei* genome (Azcarate-Peril *et al*., unpublished data). The large array of regulatory genes was proposed to be essential for rapid

response and adaptation to changing environments (9). Meanwhile, the genome of *L. acidophilus* encodes nine 2CRS, with one related to acid tolerance and proteolytic activity, and one involved in response to bile stress (20, 28, 29). Aside from RpoD, *L. acidophilus* also has an alternative sigma factor, RpoE (20). It has 96 transcriptional regulators (>5% of total genes) where majority are repressors, with only 15 transcriptional activators. Six of the transcriptional repressors are associated with sugar metabolism. Genes encoding a putative serine-threonine protein kinase and two serine-threonine protein phosphatases, all typical eukaryotic regulators, are also present in the genome. On the other hand, the transcriptional regulatory repertoires of *L. johnsonii* and *L. bulgaricus* ATCC 11842 reflect a less complex regulatory network (11, 19). In common with *L. acidophilus*, transcriptional repressors are predominant in *L. johnsonii*, with most associated with sugar uptake and metabolism via PTS. *L. bulgaricus* has only approximately one-fifth the number of putative transcription regulatory genes of *L. plantarum*. The prominently fewer number of LacI repressors in *L. bulgaricus* reflects the bacterium's evolution to the lactose-rich milk environment (11). Interestingly, the *L. bulgaricus* genome encodes tandem, cotranscribed SigA homologs that potentially arose from gene duplication. Moreover, this dairy strain has two putative extracytoplasmic function (ECF)-type sigma factors, a putative sigma factor of unknown type, and an anti-sigma factor. To date, our knowledge on the complex regulatory network involved in environmental responses in LAB is still at its infancy. Wels *et al.* (30) recently conducted a genome survey on *cis*-acting elements in *L. plantarum* using phylogenetic footprinting and identified putative regulatory motifs that are conserved among the species from the order *Bacillales* and *Lactobacillales*, as well as those unique within the LAB species. The grouping of transcriptional units with common regulatory motifs predicted regulons and reconstructed regulatory networks, which will provide a basis for gene expression analysis under specific environmental stimuli (30).

Lactobacilli frequently encounter various environmental stress conditions such as low pH, high osmotic conditions, oxidative stress, and the presence of bile during fermentation processes or transit through the GI tract. The ability to cope with these adverse conditions has been an important criterion in strain selection to ensure culture survival and performance. The following sections reviewed the *in silico* genome identification of major stress-responsive genes and functional studies that have shed light into the molecular mechanisms of stress responses in lactobacilli and the regulatory circuits involved.

4.1.2. Acid stress response and adaptation

Various mechanisms for maintaining intracellular pH homeostasis have been described in Gram-positive microorganisms, including the F_1F_0-ATPase proton pumps, amino acid decarboxylation, expression of general stress proteins and chaperones that repair or degrade damaged DNA and proteins, production of alkaline compounds, modification of cell membrane composition, and modulation of gene expression via alternative sigma factors and 2CRS (31). The *L. plantarum* genome encodes an F_1F_0-ATPase, ten sodium-proton antiporters, and three paralogous alkaline-shock proteins for maintenance of intracellular pH homeostasis (9). Using steady-state cultures, Pieterse *et al*. (32) recently showed that lactic acid specifically induced the expression of several cell surface proteins of unidentified functions that may be responsible for altering the cell surface to cope with acid stress. Among other genes that were induced, those include Clp protease, catalase, excinuclease, squalene synthase, and phytoene synthase, of which the latter two are involved in sterol biosynthesis that may increase membrane rigidity and restrict influx of lactic acid. In contrast, none of the genes coding for F_1F_0-ATPase or cation transporters were induced, prompting the authors to speculate that these components specifically respond to low pH in the presence of an inorganic acid. Likewise, a more recent investigation on the global response of acid shock in late-logarithmic phase of *L. reuteri* ATCC 55730 cells showed an induction of *clpL* encoding an ATPase with chaperone activity, genes encoding a putative esterase and a phosphatidylglycerophosphatase involved in peptidoglycan and cell membrane biosynthesis, respectively, and phage-related genes, one of which encodes a putative cell wall degradation protein (33). Also noted, the F_1F_0-ATPase operon was not differentially expressed under experimental conditions. Mutational analysis of ClpL and the esterase mutants demonstrated the importance of these genes in early response to acid shock, and possibly survival during transit through the GI tract. On the other hand, the *atp* operon encoded-F_1F_0-ATPase in *L. acidophilus* was induced when the logarithmic phase cells were subjected to acid stress at pH 3.5 with hydrochloric acid (34). In addition to the F_1F_0-ATPase, *L. acidophilus* also has several amino acid decarboxylases, a cation transport ATPase, and the chaperone Ffh that are likely involved in intracellular pH regulation (20). In particular, four genes that encode homologs involved in amino acid decarboxylation, namely, an amino acid antiporter, an amino acid permease, an ornithine decarboxylase, and a transcriptional regulator, were associated with acid tolerance in *L. acidophilus* (35). Interestingly, pre-adaptation (pH 5.5 for 1 hour) of each of all 4 acid-sensitive mutants restored the ability to survive during challenge at pH 3.5, indicating that these genes have indirect roles in acid adaptation.

4.1.3. Osmotic and oxidative stress responses

In silico genome analysis indicated that *L. plantarum* has at least three systems for the accumulation and biosynthesis of osmoprotectants glycine-betaine/carnitine/choline (9). The genome of *L. sakei* encodes three putative ABC transporters and a sodium symporter that have predicted specificities for osmo- and cyroprotectants such as betaine and carnitine, which are likely contribute to the survival in meat curing conditions (17). A trehalose utilization (*tre*) locus in the *L. acidophilus* genome encodes a PTS (*treB*) and a trehalose-6-phosphate hydrolase (*treC*) has been functionally linked to cryoprotection by both uptake and hydrolysis of trehalose (36). Interestingly, mutations in either *treB* or

treC diminished the ability of the cells to utilize trehalose for growth and serve as a cryoprotectant upon freezing and thawing cycles, suggesting that both the uptake and hydrolytic products/metabolic intermediates of trehalose contribute to its cryoprotectant effect. Perhaps one the most remarkable features of the *L. sakei* genome is its large coding capacity for enzymes involved in combating oxidative stress, a feature that is closely shared by *L. plantarum* (17). Various oxidative stress-related genes were identified in both genomes, including genes coding for catalases, peroxidases, thioredoxins, NADH oxidases, and glutathione reductases. In *L. plantarum*, it was shown that the overexpression of thioredoxin enhanced resistance against oxidative stress (37). Gene expression profiling further showed that the overexpression of thioredoxin triggered the transcription of 16 genes encoding proteins in purine metabolism, protein synthesis, stress response, and a manganese transporter that were also induced under hydrogen peroxide stress, suggesting that thioredoxin plays a role in a signal transduction cascade involved in oxidative stress response. Meanwhile, the presence of superoxide dismutase (SodA) in *L. sakei*, but not in *L. plantarum*, suggested that the *L. sakei* may be more resistant to toxic oxygen byproducts, such as superoxide and organic hydroperoxide (17). Nonetheless, the *L. plantarum* genome encodes at least five putative manganese transporters, which support previous finding that Mn^{2+} ions serve as important scavenger for oxygen radicals as well as cofactor for manganese-dependent catalase (37-39). *L. sakei* has a unique heme-dependent catalase, multiple putative iron transporters and iron-dependent regulators of the Fur family, indicating the importance of heme and iron acquisition for its adaptation to changing environmental redox conditions. A recent study using *in vivo* expression technology (IVET) also revealed the potential significance of the heat shock regulator CtsR, an L-asparaginase (that may be involved in nitrogen acquisition), and two other hypothetical proteins in the ecological adaptation and growth establishment of *L. sakei* during the initial phase of meat fermentation (40).

4.1.4. Bile stress response and tolerance

Thus far, bile salt hydrolases (BSH) or choloylglycine hydrolases have been found exclusively in human isolates, or species that encounter bile salts in their commensal habitats, such as *Lactobacillus*, *Bifidobacteria*, *L. monocytogenes*, *Enterococcus faecalis*, and *Bacteroides* (41, 42). Consistent with this observation, no BSH-encoding genes were identified in the sequenced strains of both *L. bulgaricus* and *L. helveticus*. The remaining sequenced strains possess at least one putative BSH, with the predominant number identified in *L. plantarum*, *L. johnsonii*, and *L. brevis* (Table 3). Sequence comparisons of multiple BSHs within any given strain revealed weak sequence similarity among each other. It has been speculated that the presence of multiple BSHs may provide a wider range of bile salt specificities, or exhibit differences in the mode of action depending upon bile shock or bile adaptation, which would ultimately enhance the survival of microbes in fluctuating host environment (41). For example, functional analysis of the BshA and BshB of *L. acidophilus* showed that each BSH demonstrated different substrate specificities of bile salts based on the steroid nucleus of bile salts or the amino acid side chain present in the bile salt molecules (43). Overall, the BSH orthologs in species of the *L. acidophilus* complex and *L. reuteri* are more closely related to each other than to the BSHs in other lactobacilli. On the other hand, the BSHs from *L. plantarum* (Bsh1), *L. casei* (LSEI_0412), and the pMP118-encoded BSH in *L. salivarius* (LSL_1801) shared higher sequence similarity to the BSH orthologs from *L. monocytogenes* and *Enterococcus faecium* than to other sequenced lactobacilli. In fact, no sequence homology of the BSH from *L. casei* was detected with all identified BSHs in lactobacilli at the time of analysis. These observations indicate that cross-species lateral transfer of BSH-encoding genes may be a common trend within the GI microbial community.

Although the functional characterization of the BSHs in *L. johnsonii* has not been reported, genetic analysis of the *L. johnsonii* strain 100-100 revealed two separate gene loci that encode for BSH activity (42). One of the gene loci consists of an operon that encodes two conjugated bile salt antiporters CbsT1 and CbsT2 (44) and a BSH beta peptide, while a BSH alpha peptide is encoded in another locus. Both CbsT1 and CbcT2 paralogs, which belong to the conjugated bile salt transporter (BST) family of the MFS superfamily, shared conserved sequence homology and possibly have arisen from gene duplication. Comparative analysis with a similar operon in *L. acidophilus* strain KS-13 suggested that the gene cluster was acquired via horizontal transfer due to the lack of gene synteny at the vicinity region. Recent genetic surveys of bile response in *L. acidophilus* and *L. plantarum* based on *in vitro* and *in situ* environments (29, 45, 46), have provided a broader view of the molecular mechanisms of bile tolerance in GI lactobacilli. Studies by Bron *et al.* (45, 46) using *alr* (encoding alanine racemase) complementation-based genome-wide promoter screening and microarray transcriptional analysis showed an induction of genes coding for cell envelope biosynthesis, membrane-associated proteins potentially related to bile exporter, and proteins associated with acid and oxidative stress responses in the presence of porcine bile. Intriguingly, two of the bile responsive genes were previously found to be induced during *in vivo* passage of *L. plantarum* through the mouse GI tract (47). In *L. acidophilus*, whole-genome expression analysis in response to oxgall bile revealed the upregulation of a 7-kb operon consisted of eight genes encoding a 2CRS, a transporter, a putative oxidoreductase, and four hypothetical proteins (29). Mutations within the 2CRS, the transporter, and one of the hypothetical proteins abolished bile resistance, whereas mutation in the oxidoreductase and another hypothetical protein enhanced bile tolerance, indicating a dual role of this operon in bile tolerance as well as bile sensitivity. On the other hand, mutation in the histidine protein kinase (HPK) component of the 2CRS did not affect induction of the operon in the presence of bile, whereas an increase in transcription of the operon was observed in the response regulator (RR) mutant. Therefore, this system may be part of a global regulatory network that modulates bile responsive genes. Interestingly, genes

Table 3. Degree of sequence similarity among BSH orthologs in *Lactobacillus* strains

Locus tag[1]	Protein sequence identity based on BLASTP (%)[2]																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 LBA0892 (BshA)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2 LBA1078 (BshB)	57	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3 KS-13 Bsh	33	34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4 LGAS0054	33	35	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5 LGAS0965	58	65	33	33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6 ADH Bsh	57	65	33	33	90	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7 LJ1412	57	61	38	38	56	54	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8 LJ1147	55	69	32	32	66	64	29	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9 LJ0056	33	35	92	93	33	34	37	32	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10 100-100 Cbsh-alpha	57	61	38	39	56	55	99	59	37	-	-	-	-	-	-	-	-	-	-	-	-	-
11 100-100 Cbsh-beta	33	35	93	93	33	34	37	32	99	38	-	-	-	-	-	-	-	-	-	-	-	-
12 lp_3536 (Bsh1)	52	46	37	37	47	45	52	49	38	53	38	-	-	-	-	-	-	-	-	-	-	-
13 lp_0067 (Bsh2)			30	30			30		29	30	30	32	-	-	-	-	-	-	-	-	-	-
14 lp_3362 (Bsh3)	28	29	30	30	27	27	29	25	30	28	31	30	42	-	-	-	-	-	-	-	-	-
15 lp_2572 (Bsh4)	26	26	27	27	25	24	27	26	26	27	26	28	31	26	-	-	-	-	-	-	-	-
16 Lreu0730	63	61	35	35	56	55	59	57	35	60	35	50		31		-	-	-	-	-	-	-
17 Lreu23D_0782	63	60	36	36	55	54	59	56	35	60	35	50	30	31		98	-	-	-	-	-	-
18 LSL_1801 [pMP118]	49	43	36	37	41	40	46	42	38	46	37	52	32	31	23	44	44	-	-	-	-	-
19 LSL_0518	26	26	31	31	28	26	28		32	28	32	29	32	31	25				-	-	-	-
20 LVIS_1962	28	29	29	29	28	27	30	28	30	30	29	30	56	45	31	30	30	29	29	-	-	-
21 LVIS_2132	25	25	25	25	25	24	28	25	25	28	25	27	34	30	59	26	26	23	25	33	-	-
22 LVIS_1801	25	26	25	25	27	27	26			26		25	33	29	28	26			45	32	30	-
23 LSA_0210	30	30	30	30	27	28	32	29	30	33	30		38	40	34	29	28	29	30	41	39	32
24 LSEI_0412																						

[1] LBA, *L. acidophilus* NCFM; KS-13, *L. acidophilus* KS-13, Genbank accession no. AAD03709; LGAS, *L. gasseri* ATCC 33323; ADH, *L. gasseri* ADH, Genbank accession no. AAK07837; LJ, *L. johnsonii* NCC533; 100-100, *L. johnsonii* 100-100; lp, *L. plantarum* WCFS1; Lreu, *L. reuteri* F275; Lreu23D, *L. reuteri* 100-23; LSL, *L. salivarius* UCC118; LVIS, *L. brevis* ATCC 367; LSA, *L. sakei* 23K; LSEI, *L. casei* ATCC 334 [2] number 1 to 22 correspond to the numbered putative BSHs in the first column; sequence identity ≥ 50% over the entire length of the protein sequences are shown in red; blank cells indicate no significant sequence similarity was detected between the compared sequences using BLASTP.

involved in lactose and galactose utilization and surface adherence factors were also identified as members of the bile response regulon. It was proposed that bile may act as an environmental signal and location indicator which triggered the expression of genes that are likely important for nutrient uptake from milk and intimate interaction with the host environment.

4.2. Interaction with the environments: cell surface factors

Presumably due to the importance of environmental sensing and interactions, the GI-associated lactobacilli appear to have larger and more diverse secretomes compared to the dairy-associated strains (11, 12, 48, 49). The predicted secretome of *L. plantarum* contains 233 putative extracellular proteins, of which 97 appeared unique for *L. plantarum* at the time of analysis (49). Most of these extracellular proteins are encoded in one of the two lifestyle adaptation regions of the chromosome. Comparative secretome analysis by van Pijkeren *et al*. (48) further revealed that the coding capacities for extracellular proteins are in a range of 6.2% to 9.3% among *L. salivarius*, *L. plantarum*, *L. johnsonii*, *L. acidophilus*, and *L. sakei*. In general, extracellular proteins have numerous established functions, including sensing and signaling, nutrient acquisition, and notably, interaction within the microbial population as well as with the host GI environment. Adherence factors are usually proteins or

polysaccharides that are displayed on the cell surface. Host interactions mediated by adherence factors have been associated with enhanced colonization that may promote mucosal integrity, pathogen exclusion, and host immunomodulation via host receptors recognition. In the past decades, research has been focused on understanding the interaction of commensal and probiotic lactobacilli with their host GI environment, which has led to the identification of many potential cell surface factors (50). Due to the scope of this review, the following section will focus mainly on adherence and immune factors that were identified and functionally characterized in the sequenced *Lactobacillus* strains (Figure 3).

The role of mucus-binding proteins (Mub) in intestinal mucin adherence was first described in *L. reuteri* 1063, when Roos *et al*. (51) identified a 358-kDa surface-associated Mub that selectively adhered to intestinal mucin glycoprotein. *In silico* detection of putative Mub encoded in the *Lactobacillus* genomes showed a clear bias on their predominant presence in GI-associated lactobacilli, such as *L. acidophilus* and *L. gasseri*. No Mub-encoding gene was identified in the genomes of *L. casei*, *L. bulgaricus*, *L. helveticus*, or *L. sakei* (10-12, 17). In *L. plantarum*, one of its four predicted Mubs was recently recognized as a lectin-like mannose-specific adhesin, Msa (52). The cell surface-anchored Msa protein has two mucus binding (MUB) domains similar to those of the *L. reuteri* Mub (51) and a ConA-like lectin domain, of which both domains have a predicted role in binding to mannose-containing receptors on the mucosal surface. The *L. johnsonii* genome encodes three cell wall anchored proteins that harbor MUB domains and exhibit significant sequence similarity to the Mub of *L. reuteri* 1063 (19). Interestingly, the amino acids repeat regions of two of the Mubs, LJ0047 and LJ1839, showed moderate sequence similarity to the N-terminal region of cell surface bull sperm binding protein SP18 or HR44 of *Homo sapiens*. *L. acidophilus* has thirteen putative proteins that contain one or more MUB domains, but only three that have both a predicted signal peptide sequence and a cell surface anchor LPxTG motif (R. Tallon and T. Klaenhammer, unpublished data). One of these Mub proteins, encoded by LBA1392, is composed of 4,326 residues and represents the largest protein encoded in the genome (20). Insertional inactivation of the *mub* resulted in a significant reduction in adherence to Caco-2 epithelial cells *in vitro* (53). Bioinformatic analysis of the predicted secretome in *L. salivarius* has revealed three sortase-dependent Mub homologs, Mbp-1, LspA, and LspC (48). The *mbp-1* appeared to be a pseudogene, whereas no transcription of *lspC* was detected *in vitro*. Meanwhile, a mutant of the LspA exhibited reduced ability to adhere to HT29 epithelial cells. According to van Pijkeren *et al*.(48), *lspA* may be acquired via lateral gene transfer due to its high GC content (40%) relative to the genomic GC content (33%).

Boekhorst and colleagues recently conducted bioinformatic analysis to predict and characterize MUB domain-containing proteins in LAB based on the MUB domain sequences from Mub and Msa of *L. reuteri* and *L. plantarum*, respectively (54). A total of 48 proteins with one or more MUB domains were identified, with the majority found in lactobacilli of intestinal origin. Furthermore, conserved Mub-encoding gene clusters were detected in *L. acidophilus*, *L. gasseri*, and *L. johnsonii*. Overall, the diversity of the Mubs was reflected by the domains' sequences, sizes, types, and the copy numbers. It was proposed that in a single Mub, the types of MUB domains present determine the range of adhesion targets, whereas the copy number of a similar MUB domain determines the affinity of adhesion. In addition, not all predicted Mubs possess signal peptides and a cell wall anchor LPxTG motif, indicating these proteins are either remnants or are exported via alternative mechanisms. The fact that MUB domains were found exclusively in LAB suggests that these proteins mediate specific interactions or functions between these microbes and their hosts.

A gene encoding for putative fibronectin-binding protein (Fbp) appears to be conserved in all sequenced *Lactobacillus* genomes. Interestingly, *L. brevis* also has an additional copy of *fbp* gene (LVIS_0267) in which the deduced protein shared sequence similarity to a hypothetical protein in *L. plantarum* and putative Fbp from *Enterococcus* and *Listeria* species. BlastP analysis showed that the Fbp from the *L. acidophilus* group lactobacilli are more closely related to each other (77% to 88% sequence identity) than to the Fbp in members of the other phylogenetic groups. In *L. acidophilus*, Buck *et al*. (53) showed that FbpA played a role in the attachment of Caco-2 epithelial cells, although its role in fibronectin adherence has yet to be determined. *L. plantarum* also has a large 3,360-residue surface protein, designated Sdr, with a near perfect Ser-Asp repeat of 1,600 residues (9, 49). A similar Ser-Asp repeat was found in the protein, ClfB, of *Staphylococcus aureus*, which is involved in fibrinogen adhesion (55). The presence of three putative glycosyltransferase genes at the vicinity of *clfB* has led to the speculation that the glycosyltransferases may catalyze the synthesis of mucin-like structures that coat the cell surface or mediate interaction with host cell mucins.

Major components on the cell wall, such as surface layer proteins (S-layers), techoic acids (TA), and lipotechoic acids (LTA) have also been implicated as mediators of adherence and immunomodulation. S-layers are paracrystalline arrays of proteinaceous subunits that are abundant on the cell surface of most eubacteria and archaea. It has been demonstrated that the dominant S-layer of *L. acidophilus*, encoded by *slpA*, participates in epithelial cell adherence (53). The isogenic SlpA$^-$ mutant not only has altered cell morphology, but also exhibited significantly reduced ability to bind to Caco-2 cells. The authors also considered the possibility that decreased adherence might be due to the loss of adherence factors that are typically anchored to the S-layer. Nevertheless, the SlpA of *L. brevis* ATCC 8287 has been shown to adhere to various human cell lines, fibronectin and laminin via a specific receptor-binding region (56, 57). By using monoclonal antibodies method, Granato *et al*. (58) identified LTA as the nonproteinaceous cell surface component in *L. johnsonii* that was involved in the adhesion to Caco-2 cells. A recent study also showed that

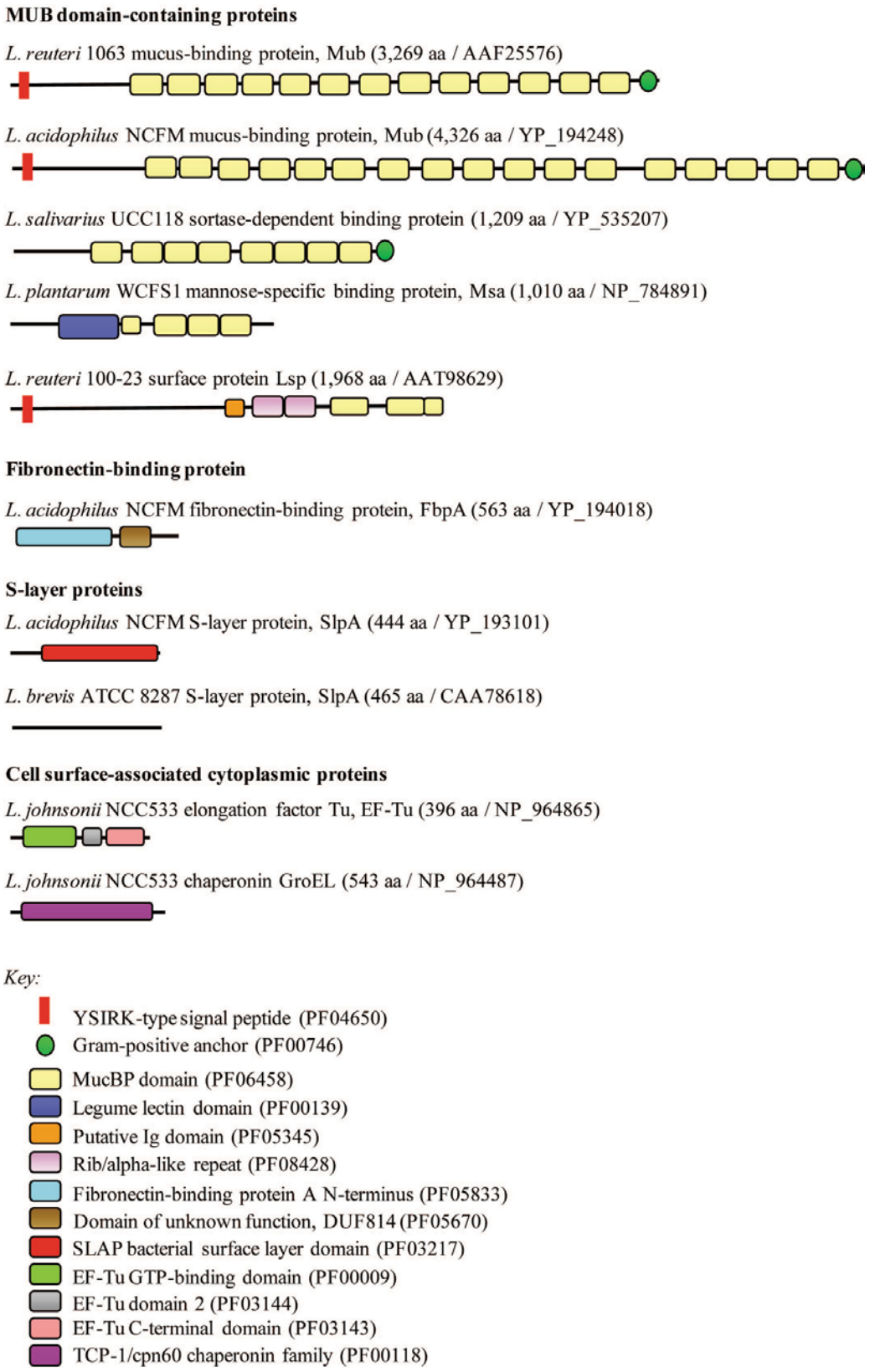

Figure 3. Adherence factors that have been functionally characterized in sequenced lactobacilli. Conserved domains and organization are depicted based on the Pfam database.

modification of the D-alanine composition in the LTA of *L. plantarum* reduced the proinflammatory response in a murine colitis model (59). Specifically, the *dltB* isogenic mutant that was defective in D-alanylation of LTA significantly enhanced the induction of interleukin (IL)-10, which in turn downregulated the inflammatory response.

Using IVET, Walter *et al*. showed that one of the genes specifically induced in *L. reuteri* 100-23 in the murine gut encodes a 300-kDa surface protein, Lsp (60, 61). Inactivation of the *lsp* gene reduced the ability of the mutant to adhere to mouse forestomach epithelium *in vivo* and *ex vivo*, as well as minimized its overall ecological performance in the murine gut. In *L. johnsonii*, it was shown that the elongation factor Tu (EF-Tu) and GroEL were involved in adherence and immunomodulation (62, 63). These studies represent the first reports describing the cell surface localization of EF-Tu and GroEL in LAB, in addition to other recent reports on the anchorless cell wall-associated DnaK, pyruvate kinase, enolase, and glyceraldehyde-3-phosphate dehydrogenase which are cytoplasmic enzymes in lactobacilli (64, 65). Both EF-Tu and GroEL preferentially adhered to intestinal epithelial cells and mucin at pH 5.0, a pH value that more closely resembles the physiological conditions in the small intestine (66). Furthermore, the recombinant EF-Tu and GroEL were able to elicit a proinflammatory response by inducing soluble CD14-dependent IL-8 secretion in HT29 epithelial cells. Recombinant GroEL expressed from *L. johnsonii* and other Gram-positive bacteria was also capable of specifically promoting aggregation of *Helicobacter pylori*. Aside from these functionally characterized cell surface factors, the genome of *L. johnsonii* also encodes a diverse repertoire of adhesion factors, some of which are unusual in LAB (19). Among the 42 predicted cell surface lipoproteins, two display weak sequence similarity to a $CD4^+$ T cell-stimulating antigen of *L. monocytogenes* and a saliva-binding protein from *S. sanguinis*, respectively. In addition, two sets of genes that collectively resembled the fimbrial operon of *Streptococcus gordonii* may encode for the biosynthesis of cell surface glycosylated fimbrial proteins similar to the Fap1 fimbrial adhesion of *Streptococcus parasanguis* and GspB platelet binding protein of *S. gordonii*. Genome analysis also revealed a putative cell surface protein that shared moderate sequence similarity to IgA proteases of pathogenic streptococci. Meanwhile, *L. johnsonii* has a fructosyltransferase homolog that was conserved among other *L. johnsonii* strains and *L. gasseri*, suggesting these species may potentially produce fructan polysaccharides that may contribute to adherence.

Despite the lack of Mub homologs in *L. sakei*, the organism has evolved with distinct sets of genes specialized for colonization on the meat matrix (17). *L. sakei* has numerous surface proteins with conserved domains that were predicted to be involved in biofilm formation, autoaggregation, and collagen adhesion. Furthermore, several identified gene clusters were predicted to encode a novel multicomponent complex on the cell surface with analogous function to the Cpf surface protein from *Lactobacillus coryneformis* that promotes coaggregation with *Escherichia coli* (67). Finally, the genome also contains two unique gene clusters that are potentially involved in the biosynthesis of surface polysaccharides, which may play a role in the adhesion to meat surfaces or the mucosa of the GI tract.

4.3. Polysaccharides biosynthetic gene clusters

Putative polysaccharide gene clusters are present in most sequenced *Lactobacillus* strains except for *L. reuteri*, *L. casei*, and *L. brevis*. The cell surface exopolysaccharide (EPS) gene clusters in the closely related *L. acidophilus*, *L. johnsonii*, and *L. gasseri* genomes are highly conserved and display high synteny with the EPS clusters in streptococci (68), although the gene organization is inverted in *L. gasseri* due to chromosomal inversion (69). All three EPS gene clusters are located within a low GC content region with a transposase gene located in the downstream region, suggesting that the *eps* clusters were acquired via horizontal gene transfer from a common source. Several EPS core genes exhibit weak or no sequence similarity between each other, despite analogous putative functions, indicating variation in the carbohydrate profiles of the EPS produced, resulting in unique surface signatures in these strains (19, 69). The EPS cluster in *L. johnsonii* encodes both secreted and cell-surface attached EPS (19). On the other hand, no EPS production was detected from *L. acidophilus* under various growth conditions examined (R. Tallon and T. Klaenhammer, unpublished data), although most of the genes within the EPS cluster were induced during log phase growth on most carbohydrates (14).

The genomes of *L. salivarius*, *L. sakei*, and *L. bulgaricus* each contains two distinct EPS gene clusters, whereas *L. plantarum* has four different surface polysaccharide gene clusters (9, 11, 13, 17). In *L. salivarius*, the EPS Cluster 1 (LSL_0977 to LSL_0997) showed little synteny with other EPS clusters, whereas Cluster 2 (LSL_1547 to LSL_1574) has an overall low GC content and displays high sequence similarity with the genes in the main surface polysaccharide clusters in *L. plantarum*, which also has noted unusual base composition (13). Recent work by Sturme *et al*. (70) showed that the expression of a surface polysaccharide locus of *L. plantarum*, *cps2* (lp_1197 to lp_1211), is regulated by the *agr*-like 2CRS involved in quorum-sensing, designated as *lam* for *Lactobacillus agr*-like module. In *L. sakei*, both surface polysaccharides gene clusters (LSA1571 to LSA1585; LSA1510 to LSA1513) encode proteins for the biosynthesis of polysaccharide-linked techoic acids and the translocation of the polysaccharides to a surface component, which contribute to adherence on meat surfaces or GI mucosa (17). In *L. bulgaricus*, both EPS clusters 1 and 2 are in close proximity to each other, with relatively conserved flanking regions (10, 11). Cluster 1 of ATCC 11842 (Ldb1937 to Ldb1957) and ATCC BAA-365 (LBUL_1800 to LBUL1815) are highly conserved in sequence and gene organization, whereas the genes in cluster 2 bear little similarity to each other in the two sequenced strains. A recent study demonstrated that the

high molecular weight acidic EPS fraction isolated from *L. delbrueckii* subsp. *bulgaricus* OLL1073R-1 was able to induce interferon-gamma production *in vitro*, and increased natural killer cell activity *in vivo* (61). Moreover, a similar immune response was observed in a mouse orally administered with yogurt fermented with OLL1073R-1 and *S. thermophilus*. Thus, it is likely that the diversity of the *eps* clusters in *L. bulgaricus* as well as in *S. thermophilus* may contribute, not only to the varied rheological properties, but also the immunostimulative characteristics of the displayed or secreted EPS.

4.4. Metabolism of unusual carbohydrates

Dietary plant polysaccharides that escape host digestion represent an important carbon source for the residing microflora in the GI tract. With the exception of *L. brevis*, *L. bulgaricus*, and *L. sakei*, the genomes of all other sequenced species encode at least one putative neopullulanase that may potentially hydrolyze pullulan, a linear polysaccharide consisting of maltotriose units linked by alpha-1,6-glucosidic bonds. Prebiotic sugars, such as fructooligosaccharide (FOS) and raffinose, have been used as non-digestible functional food ingredients to selectively promote the growth of beneficial intestinal bacteria, namely lactobacilli and bifidobacteria. Recent work on the genetic basis of FOS utilization by lactobacilli revealed diverse catabolic pathways among different species (71-73). The *L. acidophilus msm* operon encodes an ABC transporter and a cytoplasmic beta-fructofuranosidase that mediate the uptake and intracellular hydrolysis of FOS. In *L. plantarum*, FOS transport and hydrolysis involved a sucrose PTS and an intracellular beta-fructofuranosidase encoded in a sucrose gene cluster (73). *L. casei* also has a putative sugar operon similar in sequence and operon architecture with the *fos* operon in *L. paracasei* (72). FOS utilization by *L. casei* involves a putative cell-surface anchored β-fructosidase and a fructose/mannose PTS that mediate extracellular FOS degradation and uptake of fructose moieties. *L. acidophilus*, *L. plantarum*, and *L. johnsonii* were also shown to ferment raffinose (69). Whole-genome transcriptional analysis of *L. acidophilus* grown on raffinose identified a putative raffinose operon that encodes an ABC transporter and an alpha-galactosidase involved in raffinose catabolism (14).

Mucin glycoproteins serve as important growth substrate secreted by host intestinal epithelium to minimize nutrient fluctuation for the gut microbiota and thus stabilize species composition (74). The oligosaccharide side chains of epithelial mucin glycoproteins constitute of *N*-acetylgalactosamine, *N*-acetylglucosamine, *N*-acetylneuraminic acid (sialic acid), D-galactose, and L-fucose (75). Despite the apparent absence of genes encoding glycoside hydrolases required to harvest these sugar side chains, several unique genome features may enable cross-feeding of the mucin constituents generated from degradation activities by other commensals. For instance, the megaplasmid pMP118 of *L. salivarius* encodes key enzymes for scavenging of sialic acid for aminosugar metabolism (13). A putative fucose permease was identified in the genomes of *L. reuteri*, *L. plantarum*, and *L. brevis* with orthologs found in *Oenococcus oeni* and mostly Gram-negative microorganisms. Since no alpha-fucosidase gene was found in these genomes, the fucose transporter may import free fucose moieties generated from foreign hydrolysis of mucin glycoproteins. However, most intestinal lactobacilli, including *L. plantarum*, could not use fucose as sole carbon source for growth (69). Coyne *et al.* (76) recently demonstrated that *Bacteroides fragilis* accumulates host-derived L-fucose for biosynthesis of fucosylated surface capsular polysaccharides and glycoproteins via a mammalian-like pathway. Mutation of the key enzyme in this pathway rendered the mutant unable to colonize the GI tract under competitive conditions. The study revealed molecular mimicry as a novel mechanism among gut commensals that may enhance their competitiveness in the GI tract. It remains to be determined whether the fucose transporter in lactobacilli acts in a similar process, or plays other functional roles in gut survival.

4.5. Bacteriocin biosynthesis and immunity

LAB produce a diverse number of bacteriocins that, based on their structure and mode of action, are classified into Class I, II, III, or IV (77). The ability to produce bacteriocins by *Lactobacillus* is considered to provide a competitive advantage in their habitats. Practically, bacteriocin biosynthesis is a desirable characteristic for strain selection as it can be an important mechanism of pathogen exclusion in fermented foods as well as the GI environment. Using gene context analysis, Makarova *et al*. (78) recently identified putative bacteriocin gene clusters in several LAB, including *L. brevis*, *L. casei*, and *L. bulgaricus*. *L. johnsonii* has a putative operon (LJ0763b to LJ0775) that potentially encodes the production of lactacin F, a two-component class II bacteriocin (19). The putative lactacin F operon showed strong sequence similarity and synteny with the previously characterized lactacin F operon in *L. johnsonii* VPI11088 (79). The putative lactacin F operon is composed of genes encoding the bacteriocin peptide pore complex (*lafAX*), an immunity component (lafI), a 2CRS, and two ABC exporters. The presence of an IS element within the histidine kinase gene of the 2CRS, however, likely interferes with the production and regulation of lactacin F synthesis in *L. johnsonii* NCC533 (19).

Although *L. sakei* is generally regarded as common producer for sakacin P, a class IIa antilisterial bacteriocin peptide, the 23K strain does not produce sakacin P (17). Nonetheless, it has a rearranged sakacin P cluster (LSA0560_a to LSA0569_b) that contains remnants of a regulatory system and numerous genes encoding putative bacteriocin-like peptides along with cognate immunity proteins that resemble those produced by *L. plantarum* (17, 80). The presence of these immunity proteins may provide innate resistance against bacteriocins produced by closely-related competitors, which may be regarded as comparably advantageous as the ability to produce bacteriocin, since the solid nature of meat matrices can limit diffusion of antimicrobial peptides (17, 81).

Once regarded as chromosomally-encoded, the gene cluster encoding the production of the class IIb

bacteriocin, ABP-118, in *L. salivarius* was later found located on the pMP118 megaplasmid following genome sequencing (LSL_1909 to LSL_1924) (13, 82). ABP-118 is a wide spectrum salivaricin that is composed of two peptides, Abp118alpha and Abp118beta, and it is shown to inhibit species of *Bacillus*, *Listeria*, *Enterococcus*, and *Staphylococcus* (82). The production of ABP-118 is inducible by the AbpIP inducing peptide, which is part of a three-component regulatory system (3CRS) that also includes a histidine kinase and a response regulator. Meanwhile, a 9.5-kb polycistronic region that encodes the production of lactacin B (LBA1791 to LBA1803) was identified in the genome of *L. acidophilus* (20). The *lab* operon consists of 12 genes that are organized in three clusters encoding a 3CRS, a putative ABC transporter, and three unknown proteins, respectively (83). Recent functional analysis of the gene clusters has established the role of the LabT (LBA1796) as essential for the export and thus functionality of the lactacin B. Furthermore, production of lactacin B was inducible by a synthetic peptide deduced from LBA1800, indicating that this gene, along with *labK* and *labR*, encodes components of a 3CRS that presumably plays an important role in signaling and regulation of lactacin B production.

4.6. Oxalate catabolism

Oxalate is ubiquitous in dietary plant materials, and a small amount is also synthesized by the liver or generated from microbial metabolic precursors (i.e., ascorbic acid) in the GI tract (84, 85). Accumulation of oxalate at high concentrations can cause death and various pathological disorders, including hyperoxaluria, urolithiasis, pyridoxine deficiency and renal failure (86). In recent years, the role and mechanism of gut commensals in oxalate degradation and homeostasis have been recognized as important options for clinical intervention using probiotic bacteria for the treatment of oxalate-associated disorders. A recent survey of oxalate degradation activity in sixty *Lactobacillus* strains belonging to twelve species by Turroni *et al*. (87) showed that most *L. acidophilus* and *L. gasseri* strains examined exhibited high oxalate-degrading activity. Other strains that displayed moderate oxalate-degrading activity include those of *L. plantarum*, *L. casei*, and *L. rhamnosus*. In accordance with this and previous studies, the two key genes involved in oxalate catabolic pathway, *frc* (encoding formyl coenzyme A transferase) and *oxc* (oxalyl coenzyme A decarboxylase), were identified only in the sequenced strains of *L. acidophilus* (LBA0395 and LBA0396), *L. gasseri* (LGAS_0247 and LGAS_0248), *L. reuteri* F275 (Lreu0495 and Lreu0494), and *L. reuteri* 100-23 (Lreu23Draft_0637 and Lreu23Draft_0636) (87-89). Deduced amino acid sequences of both genes among the *Lactobacillus* species shared significant sequence identity (69% to 88% identity), and exhibited moderate sequence similarity with the corresponding enzymes previously characterized in *Oxalobacter formigenes* and *Bifidobacterium lactis* (88, 90, 91). Azcarate-Peril *et al.* (88) proposed that the *frc-oxc* (*oxc*) operon in NCFM may have been acquired via horizontal gene transfer based on the unusually high GC content of these genes, and the presence of the operon interrupted gene synteny in the corresponding chromosomal region in *L. johnsonii*. This observation coincides with the later findings by Turroni *et al.* (87) that, with the exception of *L. acidophilus* and *L. gasseri*, both genes and an oxalate catabolism phenotype distributed sporadically among *Lactobacillus* strains of the same species.

Global transcriptional analysis of *L. acidophilus* NCFM showed that the *frc* and *oxc* genes were induced at mildly acidic pH (pH 5.5) even in the absence of oxalate; whereas the addition of 1% ammonium oxalate at pH 6.8 resulted in down-regulation of both genes, possibly due to the dissociation of the oxalic acid ($pK_a^2 = 3.83$) (88). On the other hand, qRT-PCR analysis showed that both genes were significantly induced at pH 5.5 in the presence of 0.5% ammonium oxalate, with or without pre-adaptation to oxalate. Mutation of the *frc* gene established the role of Frc in oxalate-degrading activity in *L. acidophilus*, although a predicted oxalate permease/antiporter was not identified from *in silico* analysis and gene expression studies. In *L. gasseri*, both *frc* and *oxc* genes have been shown to co-transcribe as a single mRNA transcript (89). Like *L. acidophilus*, the *oxc* gene of *L. gasseri* was upregulated in oxalate pre-adapted cells on 0.5% sodium oxalate at pH 5.5 during mid-logarithmic growth phase. Using a three-stage continuous culture system inoculated with human fecal bacteria to simulate proximal and distal colons environment, Lewanika *et al*. (89) further demonstrated that *L. gasseri* was able to degrade oxalate in the proximal GI region, further substantiating the potential application of these *Lactobacillus* species for the management of oxalate-associated kidney diseases.

5. THE *LACTOBACILLUS* MOBILOME

The association of extrachromosomal replicons in LAB with metabolic traits such as lactose metabolism and proteolysis was first discovered in *Lactococcus lactis* more than three decades ago (92). Since then, numerous genes and operons that play important roles in the ecological adaptation of LAB have been linked to plasmids. Among the sequenced lactobacilli, a 29-kb plasmid was identified in *L. casei* (Table 2) that encodes a putative beta-glucoside utilization operon (LSEI_A04 to LSEI_A06) and an amino acid ABC transporter (LSEI_A10 to LSEI_A12), of which both gene clusters are bordered by putative transposases. *L. plantarum* also harbors a 36-kb plasmid that confers arsenate or arsenite resistance (93). The 242-kb megaplasmid pMP118, which represents the largest plasmid in LAB characterized to date, constitutes 11% of the *L. salivarius* UCC118 genome (13). Despite the absence of essential genes on pMP118, indicating its dispensability, the plasmid harbors key genes that complete the PKP, the Abp118 bacteriocin biosynthetic pathway, a bile salt hydrolase, and various carbohydrate and amino acid metabolic genes that are likely to enhance its ecological fitness (see previous sections). Recent studies showed that megaplasmids are universally present in *L. salivarius*, and were also detected in several other *Lactobacillus* species, interestingly, all of intestinal origin (13, 94). It has been proposed that the metabolic flexibility conferred by pMP118, that complements and expands

chromosomal-encoded features, will enable *L. salivarius* to cope with nutrient fluctuation in the host and adaptation to different GI niches or different hosts (13).

Genome sequencing revealed that phage DNA has a relatively substantial presence among *Lactobacillus*, which vary from polylysogenic strains to some that carry phage remnants (Table 2). All prophages identified belong to the group Sfi11-like *Siphoviridae* phage family (95-97). The majority of the prophages appear stable and noninducible under experimental conditions. Comparative analysis of prophage genomes from all sequenced lactobacilli revealed that the lysogeny modules are highly diverse and seem to follow a different evolutionary trend than the lysogens (97). Interestingly, the prophage of *L. gasseri*, LgaI, is present in tandem copies which might result from a double integration of the phage at a unique location in the chromosome (R. Barrangou and T. Klaenhammer, unpublished data). Sequence analysis of the prophages Lp1 and Lp2 in *L. plantarum* revealed the presence of several putative lysogenic conversion genes (LCG), including two mitogenic factor-like genes from a *Streptococcus pyogenes* prophage (95). This has led to the speculation that these LCG may involve in the immunomodulation of the human host. Transcriptional analysis showed that the majority of the phage genes were not expressed during lysogenic state (95-97). Transcription was detected mostly with genes near the attachment sites, including the LCG. Results from microarray and PFGE hybridizations of the sequenced species against related strains, suggested that the prophages have relatively narrow host range among lactobacilli (95, 97, 98). Genomic DNA hybridization of eight *L. johnsonii* strains against the NCC533 DNA microarray showed that a striking 50% of the NCC533 strain-specific DNA originated from its residing prophage DNA, emphasizing the contribution of prophage genome to intraspecies genetic diversity (98).

Insertion sequence (IS) elements represent another group of mobile genetic elements that have a profound impact on genetic diversity as well as genome reduction in prokaryotes (99). The genome of *L. salivarius* contains the highest number and most diversity of IS elements among the intestinal lactobacilli (13) (Table 2). Since a large proportion of these IS elements (25%) are present on pMP118, it was proposed that the formation of pMP118 may be partly mediated by IS elements via cointegration mechanisms. *L. bulgaricus* and *L. helveticus* also harbor an enormous number of intact and partial IS elements (11, 12). Despite that, IS elements played little role in the emergence of pseudogenes in *L. bulgaricus* (11). As mentioned previously, *L. helveticus* is one of the most prominent IS-loaded microorganisms sequenced to date. Unexpectedly, no extensive chromosomal rearrangement was observed in genome synteny with the *L. acidophilus* group, indicating that the presence of IS elements did not affect overall genome organization. Nonetheless, several key metabolic genes were inactivated by IS elements, notably genes encode for subunit of a beta-galactosidase and a cellobiose PTS (100, 101). Apparently IS insertion within the cellobiose *ptsC* gene was conserved across other *L. helveticus* strains examined. Cellobiose utilization is an important trait among the closely related gut lactobacilli. The loss of ability to metabolize cellobiose may indicate specialized adaptation of *L. helveticus* towards dairy environment (101).

Although no intact prophage was found in *L. acidophilus*, Altermann *et al.* (20) discovered three highly conserved potential autonomous units (PAUs) that exhibit characteristics of both plasmids and phages. The PAUs encode two restriction-modification (R/M) systems and a prophage maintenance system killer protein that may serve as a defense mechanism against phage infection. A direct repeat chromosomal region known as spacer interspersed direct repeats (SPIDR) was also identified in *L. acidophilus* and *L. helveticus* (12, 20). SPIDR represents a series of unidirectional sequence repeats alternated by nonrepetitive DNA spacers of similar size to the repeats (102). So far, no biological function has been established for SPIDR. However, SPIDR are highly polymorphic between strains, and therefore can be developed as a strain or species typing tool, as demonstrated in *Mycobacterium tuberculosis* and *Streptococcus* (20, 103, 104). The genomes of *L. helveticus*, *L. bulgaricus* and *L. casei* harbor a putative CRISPR (clustered regularly interspersed short palindromic repeats) locus. These repeat elements, complemented with *cas* gene clusters, have been implicated in phage resistance based on spacer-dependent specificity (78, 105).

6. COMPARATIVE GENOMICS

The *L. acidophilus* group represents the largest phylogenetic subgroup of *Lactobacillus*, of which many species are probiotic strains and natural inhabitants of the GI tract (1, 106). Currently, the genomes of five species from this subgroup, namely *L. acidophilus*, *L. gasseri*, *L. johnsonii*, *L. bulgaricus*, and *L. helveticus*, have been published (Table 1, Figure 1). Interspecies genome comparison showed an overall conservation in genome synteny among these species (12, 69). The intestinal associated species *L. acidophilus*, *L. johnsonii*, and *L. gasseri*, collectively known as part of the acidophilus complex, shared extensive similarity in gene content and order over the length of the genome, particularly in the latter two species, but noting a reversal of some gene order caused by genome inversion in *L. gasseri* (69, 107). As previously mentioned, all three genomes also shared a conserved low GC region that harbors EPS clusters. Microarray analysis further confirmed the close relationship between *L. johnsonii* and *L. gasseri* (107), and ORFeome comparison showed that 83-85% of the proteins were present in both genomes (19). The genome colinearity of *L. johnsonii* and *L. gasseri* was sparsely disrupted by clear insertion or deletion of gene or gene clusters into the genome backbone. Examples of *L. johnsonii*-specific gene clusters include those involved in the utilization of galactose, lactose, maltose, unusual polysaccharides, and exogenous deoxyriboses (19).

Comparative genome analyses among the dairy associated species *L. bulgaricus*, *L. helveticus*, and members of the *L. acidophilus* complex have revealed niche-specific genes that delineate this phylogenetically

close group. *L. helveticus* appeared most closely related to *L. acidophilus*, with 75% of its ORFs having orthologs in the *L. acidophilus* genome (12). Furthermore, genome alignment showed a significantly high degree of global synteny between *L. helveticus* and *L. acidophilus*, despite the presence of abundant IS elements in *L. helveticus*. Each of the two genomes contains 500 predicted genes that are not present on the other genome. Most *L. helveticus*-specific genes were clustered in two 100-kb genomic islands that encode fatty acid and amino acid metabolism, and components of R/M systems. On the other hand, *L. helveticus* showed a clear absence or inactivation of genes in *L. acidophilus* that are predicted to be important in survival and interaction within the GI tract. These include genes encoding numerous cell surface proteins, mucus-binding proteins, BSH, and sugar uptake and hydrolysis, in particular those associated with the utilization of complex dietary carbohydrates FOS and raffinose (12). One of the primary genome features that distinguishes *L. bulgaricus* ATCC 11842 is the aforementioned elevated GC content at the third codon position that has resulted in its overall genome GC content, approximately 15% higher than other members of the *L. acidophilus* group (11). About 55-60% of the ATCC 11842 proteome have orthologs in *L. acidophilus* and *L. johnsonii*, which was considered the common backbone of these three genomes, evidenced by a high conservation in gene order. In this regard, analysis of 401 core protein phylogenetic trees from these three genomes agrees with the 16S rRNA phylogeny that, despite *L. acidophilus* and *L. johnsonii* sharing a high number of orthologs, *L. acidophilus* is more closely related to ATCC 11842 than to *L. johnsonii* (108). Based on the protein tree topologies generated, Nicolas *et al*. (108) predicted that 40% of the core genes were affected by horizontal gene transfer that likely occurred between *L. acidophilus* and *L. johnsonii* which share the GI habitat. Gene content comparison based on blast analysis suggests that ATCC 11842 has a wider repertoire of peptidases, which is consistent with its dependence on peptide degradation to fulfill its amino acids requirement during growth in a milk environment. Notably, ATCC 11842 also possesses complete pathways for the biosynthesis of folate and saturated fatty acids that are not found in *L. acidophilus* and *L. johnsonii*. Like *L. helveticus*, ATCC 11842 lacks all gene features that confer potentially important probiotic traits in the *L. acidophilus* complex.

Genome comparison of *L. johnsonii* with *L. plantarum*, the latter belonging to the *L. plantarum* phylogenetic group, showed only weak conservation in whole genome synteny (107, 109). Only 28 chromosomal regions encoding nearly 550 orthologs were found conserved in gene order between the two genomes (109). Moreover, these regions, which contain genes involved in housekeeping functions, amino acids and carbohydrate metabolisms, lack colinearity, indicating major chromosomal rearrangement. Most of these gene clusters were also found in *B. subtilis*, *E. faecalis*, and *L. monocytogenes* genomes, indicating that *L. plantarum* is only marginally more related to *L. johnsonii* than to other Gram-positive species. The genome of *L. plantarum* encodes a larger number of proteins involved in the biosynthesis, uptake and metabolism of amino acids, carbohydrates, fatty acids, vitamins and cofactors compared to *L. johnsonii*, which is more dependent on exogenous nutrients. *L. plantarum* also devoted a larger proportion of its genome to regulatory genes than *L. johnsonii*. Interestingly, each genome appears to carry a unique set of cell surface proteins. Overall, analysis of inferred proteome functions reflected a rather different lifestyle between the two species. *L. plantarum* is more metabolically diverse and better equipped to thrive in various ecological niches ranging from plant environments to the human GI tract. In contrast, *L. johnsonii* is adapted to a relatively stable and nutrient rich environment in the upper GI tract (109). Meanwhile, comparative analysis of the *L. sakei* genome showed 66% conservation of orthologs with *L. plantarum*, in contrast to 42% to 43% orthologs shared with *L. johnsonii* and *L. acidophilus* (17). At the time of analysis, 63 genes were found specific for *L. sakei* that encode stress-related proteins, iron metabolism, methylglyoxal synthesis, and nucleoside catabolism, with orthologs present only in *L. monocytogenes*, *B. subtilis*, *E. faecalis*, and other aerobic Gram-positive bacteria. As discussed previously, these genes were predicted to serve important functions for *L. sakei* to promote colonization of meat environments.

A more recent genome comparison of nine sequenced species (*L. johnsonii*, *L. acidophilus*, *L. gasseri*, *L. bulgaricus* ATCC 11842, *L. casei*, *L. sakei*, *L. plantarum*, *L. salivarius*, and *L. brevis*) based on *L. johnsonii* genome as the reference further highlighted the genome diversity among the *Lactobacillus* species (107). Genome alignment at both DNA and protein levels showed highly conserved regions shared between *L. johnsonii*, *L. gasseri*, and *L. acidophilus*, with a lower degree of conservation with *L. bulgaricus*, as previously observed (11, 69, 110) (Figure 4). Despite the lack of genome sequence conservation with species outside of the *L. acidophilus* group, protein sequence similarity, in short and dispersed protein segments, could still be detected between *L. johnsonii* and these species, particularly with *L. casei* and *L. sakei* (107). On a slightly different note, comparative genome analysis of *L. acidophilus*, *L. johnsonii*, *L. plantarum*, *L. sakei*, and *L. salivarius* by Canchaya *et al*. (110) found that no single IS insertion locus or pseudogene was present in more than one genome, suggesting these events occurred after speciation and contributed to genome diversity. A total of 593 core proteins were identified for all 5 species; however, no *Lactobacillus*-specific gene was identified. Phylogenetic relationships deduced from gene synteny data suggested that unlike the 16S rRNA gene phylogeny, *L. salivarius* was more related to *L. plantarum*, whereas *L. sakei* appears as the most distantly related species among the group. A phylogenetic supertree constructed with individual trees of 354 core proteins using *E. faecalis* as an outgroup further supported the new phylogenetic position of *L. salivarius* and was comprised of four main branches: *L. acidophilus*-*L. johnsonii*, *L. salivarius*-*L. plantarum*, *L. sakei*, and *E. faecalis*. With similar whole genome approaches, the authors have proposed that the accessibility of a larger set of genome data, including other representative

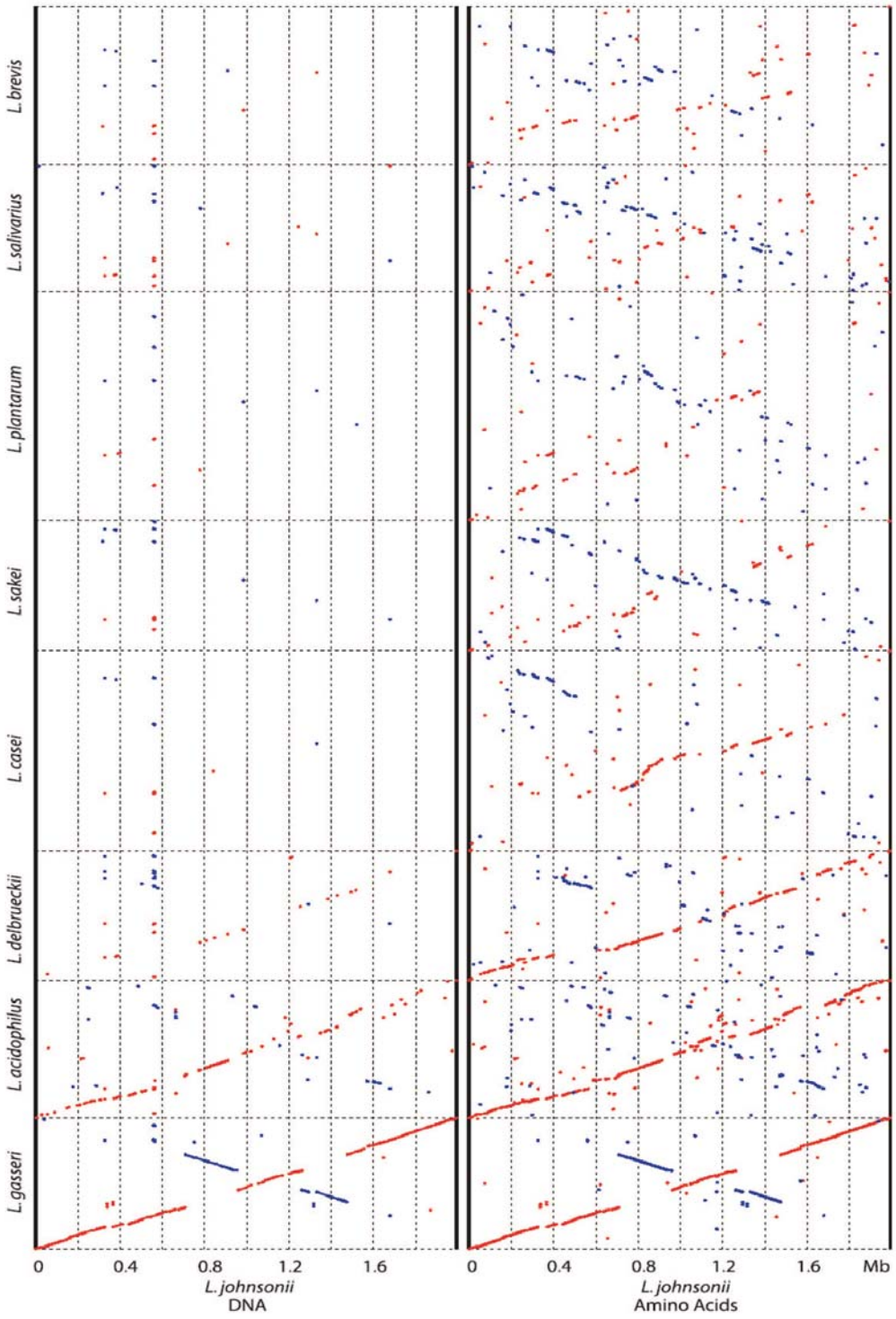

Figure 4. DNA and protein sequence similarities between completely sequenced lactobacilli (identified on the *y* axis) as revealed by *in silico* genome alignments with *L. johnsonii* NCC533, which was used as the sequenced reference strain (MUMmer analysis). (Left) Alignments obtained with NUCmer script, highlighting the conserved regions at the DNA level. The dots represent the positions of conserved DNA sequences on the genomes; (Right) Alignments obtained with PROmer script, highlighting the conserved regions at the protein level. The dots represent the positions of conserved protein sequences on the genomes. Identities in direct or reverse orientation are indicated in red and blue, respectively. Reproduced with permission from (107).

Lactobacillus species, will help define the phylogenetic relationship of this highly diverse genus.

7. GENOME EVOLUTION AND DIVERSITY

A recent study on the comprehensive analysis of nine LAB genomes proposed that niche-specific adaptation has played a central role in the evolution of LAB (10). The transition of LAB to nutritionally rich environments, such as milk and the mammalian GI tract, has resulted in parallel genome reduction and gene acquisition that would presumably enhance the ecological fitness of these microorganisms. This hypothesis is consistent with the observations that majority of the sequenced lactobacilli isolated from nutrient-rich habitats lack complete biosynthetic pathways, but have gained genes (via gene duplication or horizontal gene transfer) for acquisition of exogenous nutrient sources (11, 12, 19, 20). Based on *in silico* interpretation of the *Lactobacillus* genomes, gene decay is an ongoing trend shared among the dairy species: *L. bulgaricus*, *L. helveticus*, and the previously reported *S. thermophilus* (11, 12, 111). In *L. bulgaricus*, besides the rise of pseudogenes, the progressive reduction evolution was evidenced by a number of genome features: (i) the unusually high number of rRNA and tRNA genes (~50% higher than the average with similar genome size) that may indicate the genome has recently undergone a size reduction phase; (ii) an elevated GC content at the third codon position, which was suspected to evolve more rapidly due to its informationally neutral in translation (108) suggesting that the genome is evolving towards a higher GC content, and (iii) the presence of an inverted repeat of 47.5 kb at the replication terminus region, a conserved feature among *L. bulgaricus* strains but not in eubacteria. It was suggested that this feature represents a transient stage of genome evolution (11). Pseudogenes associated with the utilization of plant carbohydrates and the lack of complete amino acid biosynthetic pathways, was compensated for by a diverse range of amino acid and peptide transporters, peptidases, and a cell wall proteinase. These observations support a presumption that *L. bulgaricus* has shifted from a plant-associated environment to a protein-rich milk environment. Furthermore, sets of genes that provide insights into the protocooperation with *S. thermophilus* in the dairy environment were identified, including the cell wall proteinase for casein degradation, folate biosynthesis, and ornithine-putrescine exchange (11).

Reconstruction of the gene content evolution revealed that extensive loss of ancestral genes has occurred in *L. bulgaricus*, *L. johnsonii* and *L. gasseri*, whereas gene loss in *L. plantarum* and *L. casei* was counterbalanced with new and paralogous genes from horizontal gene transfer and gene duplication events (10). It is plausible that the adaptation to diverse environmental niches has exerted selective pressure to retain a majority of gene functions in *L. plantarum*, as compared to the *L. acidophilus* group members, which have evolved specialized adaptation to nutrient-stable environments. Meanwhile, comparative genomics among the closely related dairy species and the GI-associated species *L. acidophilus*, *L. gasseri*, and *L. johnsonii* further reflected the selective pressure from niche-specific adaptation on the genome evolution of these species (11, 12). The absence or decay of genes encoding carbohydrate utilization, cell surface proteins, BSH, adherence factors, bacteriocin production, and more simplified regulatory systems in *L. bulgaricus* and *L. helveticus*, presumably as a result of dairy adaptation, further highlights the importance of these genes as contributors to the probiotic functionality in these gut lactobacilli.

The plasticity of the probiotic gene repertoire among the commensal as well as transient gut lactobacilli is highly influenced by horizontal gene transfer and gene duplication events. Examples of horizontally-acquired genes include the EPS clusters in the acidophilus complex (20), various cell surface factor-encoding genes in *L. johnsonii* (19), and the lifestyle adaptation island in *L. plantarum* that encodes various carbohydrates utilization functions (9). Paralogous genes associated with carbohydrate and amino acid or peptide transporters, peptidases, mucin-binding proteins, BSHs, and extracellular proteins are also common features among the genomes of probiotic bacteria. Moreover, the presence of abundant mobile genetic elements in the genomes provides opportunities for access of niche-specific genes in the GI gene pool that may promote ecological fitness of the gut lactobacilli and overall host benefits. For instance, the LCGs recently identified in the prophage genome of *L. plantarum* were predicted to encode proteins that may play a role in host immunomodulation (see above) (95). On the other hand, several probiotic *Lactobacillus* species appear to have evolved multiple phage evasion strategies that may be important for the survival in the gut environment, noting that the GI tract represents a major reservoir for phages. For example, the previously described PAUs and CRISPR elements that potentially complement the R/M system encoded in the *L. acidophilus* genome may enhance phage resistance (20). The genome of *L. salivarius* also contains a 9.3-kb shufflon that encodes a type I R/M system and multiple HsdS subunits (13). DNA inversion-mediated phase variation of this shufflon provides nine active HsdS combinations, which was predicted to confer enhanced barrier against phage infection.

Studies devoted to examine intraspecies diversity in *Lactobacillus* have also provided insights into niche adaptive evolution within the clonal population of individual species originating from different ecological environments. In *L. plantarum*, it was found that random mutations within the arginine biosynthetic gene cluster occurred at a higher frequency in strains isolated from dairy products compared to strains isolated from plant products or humans (112). This observation suggested that arginine auxotrophy was associated with the protein-rich dairy environment. A recent study also indicated that lactose utilization is more common in *L. casei* strains isolated from cheese and human GI tract, whereas plant-associated strains were more adapted to utilize various sugar alcohols such as adonitol, ribose, sorbitol, and dulcitol (113). Furthermore, analysis of intragenic polymorphism in housekeeping genes

showed a lower nucleotide sequence diversity among *L. casei* strains isolated from the same niche, regardless of geographical regions, indicating selective pressure on the particular environmental niche. In the case of *L plantarum*, however, analysis of twenty strains (including the sequenced WCFS1 strain) derived from various ecological sources did not yield significant niche-specific genes (114). It was proposed that the metabolic diversity and versatility in the *L. plantarum* strains may enable them to survive and grow in various niches. Hence, the source of isolation may not always be a good indicator of adapted niches. In a parallel study, microarray-based genotyping comparing the sequenced WCFS1 strain with these *L. plantarum* strains revealed the variable presence of genes involved in sugar utilization, EPS biosynthesis, plantaricin biosynthesis, and NRPS biosynthesis among the strains. Two large variable regions between WCFS1 and the other strains were mapped to the lifestyle adaptation islands of the WCFS1 genome, supporting the hypothesis that these genomic islands are involved in niche adaptation. A more recent microarray-based genotyping comparing the sequenced *L. johnsonii* NCC533 strain with four other *L. johnsonii* strains has identified several regions that contribute to the diversity within this species (107). These diversity regions represent genes that are highly variable among the strains, including EPS clusters, Mub genes, a putative fimbrial biosynthesis regulon, the putative IgA protease gene, and the lactacin F gene cluster, all of which were predicted to be important in the probiotic functionality of NCC533 (19). This study addresses probiotic features as strain-specific rather than species-specific, an important consideration in terms of probiotic candidate selection (107).

8. FUTURE PERSPECTIVE

The growing number of *Lactobacillus* genome sequences has undoubtedly served as a springboard for comparative analysis and functional characterization of genotypic traits that are important for bioproccessing and probiotic functionality in lactobacilli. In terms of probiotic development, ongoing targeted functional analyses in probiotic strains have continued to revolutionize our view on the mechanisms of environmental adaptation and host interaction in the GI environment. The accessibility of complete genomes also facilitates the development of high-throughput functional genomic tools for studies on complex regulatory networks that govern probiotic activities *in vivo*. Development of DNA microarray platforms integrated with polyphasic analysis allows evaluation of strain-specific probiotic traits that will aid in strain selection. Moreover, genome sequencing of autochthonous and allochthonous *Lactobacillus* population of the human GI tract will provide insights into the pan-genome and genome plasticity of a given probiotic species. The resulting broader view of the genetic landscape will lead to the identification of unique or desirable gene features that will aid in strain improvement and tailored biotherapeutics. Meanwhile, the collective genomes of the gut microbiota, termed "microbiome", were predicted to encode at least 100 times the number of genes in the human genome (74), some of which complement the physiological and metabolic functions in their host. Regarded as part of the gut microbiome, the genome data of probiotic lactobacilli will contribute to the metagenomic analysis of the gut microbiota, which will ultimately provide mechanistic insights into the intertwined relationships between the GI community and their human hosts.

9. ACKNOWLEDGMENTS

The research program on probiotic lactobacilli at North Carolina State University is supported by the North Carolina Dairy Foundation, Danisco USA, Inc., Dairy Management Inc., and the Southeast Dairy Foods Research Center.

10. REFERENCES

1. Felis, G. E. and F. Dellaglio: Taxonomy of lactobacilli and bifidobacteria. *Curr Issues Intest Microbiol*, 8, 44-61 (2007)

2. Walter, J.: The microecology of lactobacilli in the gastrointestinal tract. In: Probiotics and Prebiotics: scientific aspects. Ed: G. W. Tannock. Caister Academic Press, Norfolk, UK (2005)

3. Stiles, M. E.: Biopreservation by lactic acid bacteria. *Antonie Van Leeuwenhoek*, 70, 331-345 (1996)

4. Reid, G., M. E. Sanders, H. R. Gaskins, G. R. Gibson, A. Mercenier, R. Rastall, M. Roberfroid, I. Rowland, C. Cherbut and T. R. Klaenhammer: New scientific paradigms for probiotics and prebiotics. *J Clin Gastroenterol*, 37, 105-118 (2003)

5. Yan, F., H. Cao, T. L. Cover, R. Whitehead, M. K. Washington and D. B. Polk: Soluble proteins produced by probiotic bacteria regulate intestinal epithelial cell survival and growth. *Gastroenterology*, 132, 562-575 (2007)

6. Chen, X., J. Xu, J. Shuai, J. Chen, Z. Zhang and W. Fang: The S-layer proteins of *Lactobacillus crispatus* strain ZJ001 is responsible for competitive exclusion against *Escherichia coli* O157:H7 and *Salmonella typhimurium*. *Int J Food Microbiol*, 115, 307-312 (2007)

7. Foligne, B., G. Zoumpopoulou, J. Dewulf, A. Ben Younes, F. Chareyre, J. C. Sirard, B. Pot and C. Grangette: A key role of dendritic cells in probiotic functionality. *PLoS ONE*, 2, e313 (2007)

8. Mohamadzadeh, M., S. Olson, W. V. Kalina, G. Ruthel, G. L. Demmin, K. L. Warfield, S. Bavari and T. R. Klaenhammer: Lactobacilli activate human dendritic cells that skew T cells toward T helper 1 polarization. *Proc Natl Acad Sci U S A*, 102, 2880-2885 (2005)

9. Kleerebezem, M., J. Boekhorst, R. van Kranenburg, D. Molenaar, O. P. Kuipers, R. Leer, R. Tarchini, S. A. Peters, H. M. Sandbrink, M. W. Fiers, W. Stiekema, R. M. Lankhorst, P. A. Bron, S. M. Hoffer, M. N. Groot, R. Kerkhoven, M. de Vries, B. Ursing, W. M. de Vos and R. J. Siezen: Complete genome sequence of *Lactobacillus*

plantarum WCFS1. *Proc Natl Acad Sci U S A*, 100, 1990-1995 (2003)

10. Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, B. Mirkin, E. Koonin, A. Pavlov, N. Pavlova, V. Karamychev, N. Polouchine, V. Shakhova, I. Grigoriev, Y. Lou, D. Rohksar, S. Lucas, K. Huang, D. M. Goodstein, T. Hawkins, V. Plengvidhya, D. Welker, J. Hughes, Y. Goh, A. Benson, K. Baldwin, J. H. Lee, I. Diaz-Muniz, B. Dosti, V. Smeianov, W. Wechter, R. Barabote, G. Lorca, E. Altermann, R. Barrangou, B. Ganesan, Y. Xie, H. Rawsthorne, D. Tamir, C. Parker, F. Breidt, J. Broadbent, R. Hutkins, D. O'Sullivan, J. Steele, G. Unlu, M. Saier, T. Klaenhammer, P. Richardson, S. Kozyavkin, B. Weimer and D. Mills: Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A*, 103, 15611-15616 (2006)

11. van de Guchte, M., S. Penaud, C. Grimaldi, V. Barbe, K. Bryson, P. Nicolas, C. Robert, S. Oztas, S. Mangenot, A. Couloux, V. Loux, R. Dervyn, R. Bossy, A. Bolotin, J. M. Batto, T. Walunas, J. F. Gibrat, P. Bessieres, J. Weissenbach, S. D. Ehrlich and E. Maguin: The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. *Proc Natl Acad Sci U S A*, 103, 9274-9279 (2006)

12. Callanan, M., P. Kaleta, J. O'Callaghan, O. O'Sullivan, K. Jordan, O. McAuliffe, A. Sangrador-Vegas, L. Slattery, G. F. Fitzgerald, T. Beresford and R. P. Ross: Genome sequence of *Lactobacillus helveticus*, an organism distinguished by selective gene loss and insertion sequence element expansion. *J Bacteriol*, 190, 727-735 (2008)

13. Claesson, M. J., Y. Li, S. Leahy, C. Canchaya, J. P. van Pijkeren, A. M. Cerdeno-Tarraga, J. Parkhill, S. Flynn, G. C. O'Sullivan, J. K. Collins, D. Higgins, F. Shanahan, G. F. Fitzgerald, D. van Sinderen and P. W. O'Toole: Multireplicon genome architecture of *Lactobacillus salivarius*. *Proc Natl Acad Sci U S A*, 103, 6718-6723 (2006)

14. Barrangou, R., M. A. Azcarate-Peril, T. Duong, S. B. Conners, R. M. Kelly and T. R. Klaenhammer: Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays. *Proc Natl Acad Sci U S A*, 103, 3816-3821 (2006)

15. Arskold, E., E. Lohmeier-Vogel, R. Cao, S. Roos, P. Radstrom and E. W. van Niel: Phosphoketolase pathway dominates in *Lactobacillus reuteri* ATCC 55730 containing dual pathways for glycolysis. *J Bacteriol*, 190, 206-212 (2008)

16. Li, Y., E. Raftis, C. Canchaya, G. F. Fitzgerald, D. van Sinderen and P. W. O'Toole: Polyphasic analysis indicates that *Lactobacillus salivarius* subsp. *salivarius* and *Lactobacillus salivarius* subsp. *salicinius* do not merit separate subspecies status. *Int J Syst Evol Microbiol*, 56, 2397-2403 (2006)

17. Chaillou, S., M. C. Champomier-Verges, M. Cornet, A. M. Crutz-Le Coq, A. M. Dudez, V. Martin, S. Beaufils, E. Darbon-Rongere, R. Bossy, V. Loux and M. Zagorec: The complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K. *Nat Biotechnol*, 23, 1527-1533 (2005)

18. Taranto, M. P., J. L. Vera, J. Hugenholtz, G. F. De Valdez and F. Sesma: *Lactobacillus reuteri* CRL1098 produces cobalamin. *J Bacteriol*, 185, 5643-5647 (2003)

19. Pridmore, R. D., B. Berger, F. Desiere, D. Vilanova, C. Barretto, A. C. Pittet, M. C. Zwahlen, M. Rouvet, E. Altermann, R. Barrangou, B. Mollet, A. Mercenier, T. Klaenhammer, F. Arigoni and M. A. Schell: The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. *Proc Natl Acad Sci U S A*, 101, 2512-2517 (2004)

20. Altermann, E., W. M. Russell, M. A. Azcarate-Peril, R. Barrangou, B. L. Buck, O. McAuliffe, N. Souther, A. Dobson, T. Duong, M. Callanan, S. Lick, A. Hamrick, R. Cano and T. R. Klaenhammer: Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc Natl Acad Sci U S A*, 102, 3906-3912 (2005)

21. Hols, P., F. Hancy, L. Fontaine, B. Grossiord, D. Prozzi, N. Leblond-Bourget, B. Decaris, A. Bolotin, C. Delorme, S. Dusko Ehrlich, E. Guedon, V. Monnet, P. Renault and M. Kleerebezem: New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. *FEMS Microbiol Rev*, 29, 435-463 (2005)

22. Siezen, R., J. Boekhorst, L. Muscariello, D. Molenaar, B. Renckens and M. Kleerebezem: *Lactobacillus plantarum* gene clusters encoding putative cell-surface protein complexes for carbohydrate utilization are conserved in specific gram-positive bacteria. *BMC Genomics*, 7, 126 (2006)

23. Welman, A. D. and I. S. Maddox: Fermentation performance of an exopolysaccharide-producing strain of *Lactobacillus delbrueckii* subsp. *bulgaricus*. *J Ind Microbiol Biotechnol*, 30, 661-668 (2003)

24. Romano, A. H., J. D. Trifone and M. Brustolon: Distribution of the phosphoenolpyruvate:glucose phosphotransferase system in fermentative bacteria. *J Bacteriol*, 139, 93-97 (1979)

25. Saier, M. H., Jr., J. J. Ye, S. Klinke and E. Nino: Identification of an anaerobically induced phosphoenolpyruvate-dependent fructose-specific phosphotransferase system and evidence for the Embden-Meyerhof glycolytic pathway in the heterofermentative bacterium *Lactobacillus brevis*. *J Bacteriol*, 178, 314-316 (1996)

26. Taranto, M. P., G. Font de Valdez and G. Perez-Martinez: Evidence of a glucose proton motive force-

dependent permease and a fructose phosphoenolpyruvate:phosphotransferase transport system in *Lactobacillus reuteri* CRL 1098. *FEMS Microbiol Lett*, 181, 109-112 (1999)

27. Kornberg, H. L. and J. Smith: Role of phosphofructokinase in the utilization of glucose by *Escherichia coli*. *Nature*, 227, 44-46 (1970)

28. Azcarate-Peril, M. A., O. McAuliffe, E. Altermann, S. Lick, W. M. Russell and T. R. Klaenhammer: Microarray analysis of a two-component regulatory system involved in acid resistance and proteolytic activity in *Lactobacillus acidophilus*. *Appl Environ Microbiol*, 71, 5794-5804 (2005)

29. Pfeiler, E. A., M. A. Azcarate-Peril and T. R. Klaenhammer: Characterization of a novel bile-inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus*. *J Bacteriol*, 189, 4624-4634 (2007)

30. Wels, M., C. Francke, R. Kerkhoven, M. Kleerebezem and R. J. Siezen: Predicting cis-acting elements of *Lactobacillus plantarum* by comparative genomics with different taxonomic subgroups. *Nucleic Acids Res*, 34, 1947-1958 (2006)

31. Cotter, P. D. and C. Hill: Surviving the acid test: responses of Gram-positive bacteria to low pH. *Microbiol Mol Biol Rev*, 67, 429-453 (2003)

32. Pieterse, B., R. J. Leer, F. H. Schuren and M. J. van der Werf: Unravelling the multiple effects of lactic acid stress on *Lactobacillus plantarum* by transcription profiling. *Microbiology*, 151, 3881-3894 (2005)

33. Wall, T., K. Bath, R. A. Britton, H. Jonsson, J. Versalovic and S. Roos: The early response to acid shock in *Lactobacillus reuteri* involves the ClpL chaperone and a putative cell wall-altering esterase. *Appl Environ Microbiol*, 73, 3924-3935 (2007)

34. Kullen, M. J. and T. R. Klaenhammer: Identification of the pH-inducible, proton-translocating F_1F_0-ATPase (*atpBEFHAGDC*) operon of *Lactobacillus acidophilus* by differential display: gene structure, cloning and characterization. *Mol Microbiol*, 33, 1152-1161 (1999)

35. Azcarate-Peril, M. A., E. Altermann, R. L. Hoover-Fitzula, R. J. Cano and T. R. Klaenhammer: Identification and inactivation of genetic loci involved with *Lactobacillus acidophilus* acid tolerance. *Appl Environ Microbiol*, 70, 5315-5322 (2004)

36. Duong, T., R. Barrangou, W. M. Russell and T. R. Klaenhammer: Characterization of the tre locus and analysis of trehalose cryoprotection in *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol*, 72, 1218-1225 (2006)

37. Serrano, L. M., D. Molenaar, M. Wels, B. Teusink, P. A. Bron, W. M. de Vos and E. J. Smid: Thioredoxin reductase is a key factor in the oxidative stress response of *Lactobacillus plantarum* WCFS1. *Microb Cell Fact*, 6, 29 (2007)

38. Groot, M. N., E. Klaassens, W. M. de Vos, J. Delcour, P. Hols and M. Kleerebezem: Genome-based in silico detection of putative manganese transport systems in *Lactobacillus plantarum* and their genetic analysis. *Microbiology*, 151, 1229-1238 (2005)

39. Archibald, F. S. and I. Fridovich: Manganese and defenses against oxygen toxicity in *Lactobacillus plantarum*. *J Bacteriol*, 145, 442-451 (1981)

40. Hufner, E., T. Markieton, S. Chaillou, A. M. Crutz-Le Coq, M. Zagorec and C. Hertel: Identification of *Lactobacillus sakei* genes induced during meat fermentation and their role in survival and growth. *Appl Environ Microbiol*, 73, 2522-2531 (2007)

41. Begley, M., C. Hill and C. G. Gahan: Bile salt hydrolase activity in probiotics. *Appl Environ Microbiol*, 72, 1729-1738 (2006)

42. Elkins, C. A., S. A. Moser and D. C. Savage: Genes encoding bile salt hydrolases and conjugated bile salt transporters in *Lactobacillus johnsonii* 100-100 and other *Lactobacillus* species. *Microbiology*, 147, 3403-3412 (2001)

43. McAuliffe, O., R. J. Cano and T. R. Klaenhammer: Genetic analysis of two bile salt hydrolase activities in *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol*, 71, 4925-4929 (2005)

44. Elkins, C. A. and D. C. Savage: CbsT2 from *Lactobacillus johnsonii* 100-100 is a transport protein of the major facilitator superfamily that facilitates bile acid antiport. *J Mol Microbiol Biotechnol*, 6, 76-87 (2003)

45. Bron, P. A., M. Marco, S. M. Hoffer, E. Van Mullekom, W. M. de Vos and M. Kleerebezem: Genetic characterization of the bile salt response in *Lactobacillus plantarum* and analysis of responsive promoters *in vitro* and *in situ* in the gastrointestinal tract. *J Bacteriol*, 186, 7829-7835 (2004)

46. Bron, P. A., D. Molenaar, W. M. de Vos and M. Kleerebezem: DNA micro-array-based identification of bile-responsive genes in *Lactobacillus plantarum*. *J Appl Microbiol*, 100, 728-738 (2006)

47. Bron, P. A., C. Grangette, A. Mercenier, W. M. de Vos and M. Kleerebezem: Identification of *Lactobacillus plantarum* genes that are induced in the gastrointestinal tract of mice. *J Bacteriol*, 186, 5721-5729 (2004)

48. van Pijkeren, J. P., C. Canchaya, K. A. Ryan, Y. Li, M. J. Claesson, B. Sheil, L. Steidler, L. O'Mahony, G. F. Fitzgerald, D. van Sinderen and P. W. O'Toole:

Comparative and functional analysis of sortase-dependent proteins in the predicted secretome of *Lactobacillus salivarius* UCC118. *Appl Environ Microbiol*, 72, 4143-4153 (2006)

49. Boekhorst, J., M. Wels, M. Kleerebezem and R. J. Siezen: The predicted secretome of *Lactobacillus plantarum* WCFS1 sheds light on interactions with its environment. *Microbiology*, 152, 3175-3183 (2006)

50. Velez, M. P., S. C. De Keersmaecker and J. Vanderleyden: Adherence factors of *Lactobacillus* in the human gastrointestinal tract. *FEMS Microbiol Lett*, 276, 140-148 (2007)

51. Roos, S. and H. Jonsson: A high-molecular-mass cell-surface protein from *Lactobacillus reuteri* 1063 adheres to mucus components. *Microbiology*, 148, 433-442 (2002)

52. Pretzer, G., J. Snel, D. Molenaar, A. Wiersma, P. A. Bron, J. Lambert, W. M. de Vos, R. van der Meer, M. A. Smits and M. Kleerebezem: Biodiversity-based identification and functional characterization of the mannose-specific adhesin of *Lactobacillus plantarum*. *J Bacteriol*, 187, 6128-6136 (2005)

53. Buck, B. L., E. Altermann, T. Svingerud and T. R. Klaenhammer: Functional analysis of putative adhesion factors in *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol*, 71, 8344-8351 (2005)

54. Boekhorst, J., Q. Helmer, M. Kleerebezem and R. J. Siezen: Comparative analysis of proteins with a mucus-binding domain found exclusively in lactic acid bacteria. *Microbiology*, 152, 273-280 (2006)

55. Hartford, O., P. Francois, P. Vaudaux and T. J. Foster: The dipeptide repeat region of the fibrinogen-binding protein (clumping factor) is required for functional expression of the fibrinogen-binding domain on the *Staphylococcus aureus* cell surface. *Mol Microbiol*, 25, 1065-1076 (1997)

56. de Leeuw, E., X. Li and W. Lu: Binding characteristics of the *Lactobacillus brevis* ATCC 8287 surface layer to extracellular matrix proteins. *FEMS Microbiol Lett*, 260, 210-215 (2006)

57. Hynonen, U., B. Westerlund-Wikstrom, A. Palva and T. K. Korhonen: Identification by flagellum display of an epithelial cell- and fibronectin-binding function in the SlpA surface protein of *Lactobacillus brevis*. *J Bacteriol*, 184, 3360-3367 (2002)

58. Granato, D., F. Perotti, I. Masserey, M. Rouvet, M. Golliard, A. Servin and D. Brassart: Cell surface-associated lipoteichoic acid acts as an adhesion factor for attachment of *Lactobacillus johnsonii* La1 to human enterocyte-like Caco-2 cells. *Appl Environ Microbiol*, 65, 1071-1077 (1999)

59. Grangette, C., S. Nutten, E. Palumbo, S. Morath, C. Hermann, J. Dewulf, B. Pot, T. Hartung, P. Hols and A. Mercenier: Enhanced antiinflammatory capacity of a *Lactobacillus plantarum* mutant synthesizing modified teichoic acids. *Proc Natl Acad Sci U S A*, 102, 10321-10326 (2005)

60. Walter, J., N. C. Heng, W. P. Hammes, D. M. Loach, G. W. Tannock and C. Hertel: Identification of *Lactobacillus reuteri* genes specifically induced in the mouse gastrointestinal tract. *Appl Environ Microbiol*, 69, 2044-2051 (2003)

61. Walter, J., P. Chagnaud, G. W. Tannock, D. M. Loach, F. Dal Bello, H. F. Jenkinson, W. P. Hammes and C. Hertel: A high-molecular-mass surface protein (Lsp) and methionine sulfoxide reductase B (MsrB) contribute to the ecological performance of *Lactobacillus reuteri* in the murine gut. *Appl Environ Microbiol*, 71, 979-986 (2005)

62. Granato, D., G. E. Bergonzelli, R. D. Pridmore, L. Marvin, M. Rouvet and I. E. Corthesy-Theulaz: Cell surface-associated elongation factor Tu mediates the attachment of *Lactobacillus johnsonii* NCC533 (La1) to human intestinal cells and mucins. *Infect Immun*, 72, 2160-2169 (2004)

63. Bergonzelli, G. E., D. Granato, R. D. Pridmore, L. F. Marvin-Guy, D. Donnicola and I. E. Corthesy-Theulaz: GroEL of *Lactobacillus johnsonii* La1 (NCC 533) is cell surface associated: potential role in interactions with the host and the gastric pathogen *Helicobacter pylori*. *Infect Immun*, 74, 425-434 (2006)

64. Kelly, P., P. B. Maguire, M. Bennett, D. J. Fitzgerald, R. J. Edwards, B. Thiede, A. Treumann, J. K. Collins, G. C. O'Sullivan, F. Shanahan and C. Dunne: Correlation of probiotic *Lactobacillus salivarius* growth phase with its cell wall-associated proteome. *FEMS Microbiol Lett*, 252, 153-159 (2005)

65. Hurmalainen, V., S. Edelman, J. Antikainen, M. Baumann, K. Lahteenmaki and T. K. Korhonen: Extracellular proteins of *Lactobacillus crispatus* enhance activation of human plasminogen. *Microbiology*, 153, 1112-1122 (2007)

66. Blum, S., R. Reniero, E. J. Schiffrin, R. Crittenden, T. Mattila-Sandholm, A. C. Ouwehand, S. Salminen, A. von Wright, M. Saarela, M. Saxelin, K. Collins and L. Morelli: Adhesion studies for probiotics: need for validation and refinement. *Trends Food Sci Technol*, 10, 405-410 (2000)

67. Schachtsiek, M., W. P. Hammes and C. Hertel: Characterization of *Lactobacillus coryniformis* DSM 20001^T surface protein Cpf mediating coaggregation with and aggregation among pathogens. *Appl Environ Microbiol*, 70, 7078-7085 (2004)

68. Stingele, F., J. R. Neeser and B. Mollet: Identification and characterization of the *eps* (exopolysaccharide) gene cluster from *Streptococcus thermophilus* Sfi6. *J Bacteriol*, 178, 1680-1690 (1996)

69. Klaenhammer, T. R., R. Barrangou, B. L. Buck, M. A. Azcarate-Peril and E. Altermann: Genomic features of

lactic acid bacteria effecting bioprocessing and health. *FEMS Microbiol Rev*, 29, 393-409 (2005)

70. Sturme, M. H., J. Nakayama, D. Molenaar, Y. Murakami, R. Kunugi, T. Fujii, E. E. Vaughan, M. Kleerebezem and W. M. de Vos: An *agr*-like two-component regulatory system in *Lactobacillus plantarum* is involved in production of a novel cyclic peptide and regulation of adherence. *J Bacteriol*, 187, 5224-5235 (2005)

71. Barrangou, R., E. Altermann, R. Hutkins, R. Cano and T. R. Klaenhammer: Functional and comparative genomic analyses of an operon involved in fructooligosaccharide utilization by *Lactobacillus acidophilus*. *Proc Natl Acad Sci U S A*, 100, 8957-8962 (2003)

72. Goh, Y. J., C. Zhang, A. K. Benson, V. Schlegel, J. H. Lee and R. W. Hutkins: Identification of a putative operon involved in fructooligosaccharide utilization by *Lactobacillus paracasei*. *Appl Environ Microbiol*, 72, 7518-7530 (2006)

73. Saulnier, D. M., D. Molenaar, W. M. de Vos, G. R. Gibson and S. Kolida: Identification of prebiotic fructooligosaccharide metabolism in *Lactobacillus plantarum* WCFS1 through microarrays. *Appl Environ Microbiol*, 73, 1753-1765 (2007)

74. Backhed, F., R. E. Ley, J. L. Sonnenburg, D. A. Peterson and J. I. Gordon: Host-bacterial mutualism in the human intestine. *Science*, 307, 1915-1920 (2005)

75. Murty, V. L., F. J. Downs and W. Pigman: Rat-colonic, mucus glycoprotein. *Carbohydr Res*, 61, 139-145 (1978)

76. Coyne, M. J., B. Reinap, M. M. Lee and L. E. Comstock: Human symbionts use a host-like pathway for surface fucosylation. *Science*, 307, 1778-1781 (2005)

77. Klaenhammer, T. R.: Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiol Rev*, 12, 39-85 (1993)

78. Makarova, K. S. and E. V. Koonin: Evolutionary genomics of lactic acid bacteria. *J Bacteriol*, 189, 1199-1208 (2007)

79. Allison, G. E., C. Fremaux and T. R. Klaenhammer: Expansion of bacteriocin activity and host range upon complementation of two peptides encoded within the lactacin F operon. *J Bacteriol*, 176, 2235-2241 (1994)

80. Moretro, T., K. Naterstad, E. Wang, I. M. Aasen, S. Chaillou, M. Zagorec and L. Axelsson: Sakacin P non-producing Lactobacillus sakei strains contain homologues of the sakacin P gene cluster. Res Microbiol, 156, 949-960 (2005)

81. Vermeiren, L., F. Devlieghere and J. Debevere: Evaluation of meat born lactic acid bacteria as protective cultures for the biopreservation of cooked meat products. Int J Food Microbiol, 96, 149-164 (2004)

82. Flynn, S., D. van Sinderen, G. M. Thornton, H. Holo, I. F. Nes and J. K. Collins: Characterization of the genetic locus responsible for the production of ABP-118, a novel bacteriocin produced by the probiotic bacterium *Lactobacillus salivarius* subsp. *salivarius* UCC118. *Microbiology*, 148, 973-984 (2002)

83. Dobson, A. E., R. B. Sanozky-Dawes and T. R. Klaenhammer: Identification of an operon and inducing peptide involved in the production of lactacin B by *Lactobacillus acidophilus*. *J Appl Microbiol*, 103, 1766-1778 (2007)

84. Holmes, R. P. and D. G. Assimos: Glyoxylate synthesis, and its modulation and influence on oxalate synthesis. *J Urol*, 160, 1617-1624 (1998)

85. Ogawa, Y., T. Miyazato and T. Hatano: Oxalate and urinary stones. *World J Surg*, 24, 1154-1159 (2000)

86. Hatch, M. and R. W. Freel: Alterations in intestinal transport of oxalate in disease states. *Scanning Microsc*, 9, 1121-1126 (1995)

87. Turroni, S., B. Vitali, C. Bendazzoli, M. Candela, R. Gotti, F. Federici, F. Pirovano and P. Brigidi: Oxalate consumption by lactobacilli: evaluation of oxalyl-CoA decarboxylase and formyl-CoA transferase activity in *Lactobacillus acidophilus*. *J Appl Microbiol*, 103, 1600-1609 (2007)

88. Azcarate-Peril, M. A., J. M. Bruno-Barcena, H. M. Hassan and T. R. Klaenhammer: Transcriptional and functional analysis of oxalyl-coenzyme A (CoA) decarboxylase and formyl-CoA transferase genes from *Lactobacillus acidophilus*. *Appl Environ Microbiol*, 72, 1891-1899 (2006)

89. Lewanika, T. R., S. J. Reid, V. R. Abratt, G. T. Macfarlane and S. Macfarlane: *Lactobacillus gasseri* Gasser AM63(T) degrades oxalate in a multistage continuous culture simulator of the human colonic microbiota. *FEMS Microbiol Ecol*, 61, 110-120 (2007)

90. Lung, H. Y., A. L. Baetz and A. B. Peck: Molecular cloning, DNA sequence, and gene expression of the oxalyl-coenzyme A decarboxylase gene, oxc, from the bacterium *Oxalobacter formigenes*. *J Bacteriol*, 176, 2468-2472 (1994)

91. Federici, F., B. Vitali, R. Gotti, M. R. Pasca, S. Gobbi, A. B. Peck and P. Brigidi: Characterization and heterologous expression of the oxalyl coenzyme A decarboxylase gene from *Bifidobacterium lactis*. *Appl Environ Microbiol*, 70, 5066-5073 (2004)

92. Cords, B. R., L. L. McKay and P. Guerry: Extrachromosomal elements in group N streptococci. *J Bacteriol*, 117, 1149-1152 (1974)

93. van Kranenburg, R., N. Golic, R. Bongers, R. J. Leer, W. M. de Vos, R. J. Siezen and M. Kleerebezem:

Functional analysis of three plasmids from *Lactobacillus plantarum*. *Appl Environ Microbiol*, 71, 1223-1230 (2005)

94. Li, Y., C. Canchaya, F. Fang, E. Raftis, K. A. Ryan, J. P. van Pijkeren, D. van Sinderen and P. W. O'Toole: Distribution of megaplasmids in *Lactobacillus salivarius* and other lactobacilli. *J Bacteriol*, 189, 6128-6139 (2007)

95. Ventura, M., C. Canchaya, M. Kleerebezem, W. M. de Vos, R. J. Siezen and H. Brussow: The prophage sequences of *Lactobacillus plantarum* strain WCFS1. *Virology*, 316, 245-255 (2003)

96. Ventura, M., C. Canchaya, R. D. Pridmore and H. Brussow: The prophages of *Lactobacillus johnsonii* NCC 533: comparative genomics and transcription analysis. *Virology*, 320, 229-242 (2004)

97. Ventura, M., C. Canchaya, V. Bernini, E. Altermann, R. Barrangou, S. McGrath, M. J. Claesson, Y. Li, S. Leahy, C. D. Walker, R. Zink, E. Neviani, J. Steele, J. Broadbent, T. R. Klaenhammer, G. F. Fitzgerald, W. O'Toole P and D. van Sinderen: 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 (2006)

98. Ventura, M., C. Canchaya, D. Pridmore, B. Berger and H. Brussow: Integration and distribution of *Lactobacillus johnsonii* prophages. *J Bacteriol*, 185, 4603-4608 (2003)

99. Parkhill, J., M. Sebaihia, A. Preston, L. D. Murphy, N. Thomson, D. E. Harris, M. T. Holden, C. M. Churcher, S. D. Bentley, K. L. Mungall, A. M. Cerdeno-Tarraga, L. Temple, K. James, B. Harris, M. A. Quail, M. Achtman, R. Atkin, S. Baker, D. Basham, N. Bason, I. Cherevach, T. Chillingworth, M. Collins, A. Cronin, P. Davis, J. Doggett, T. Feltwell, A. Goble, N. Hamlin, H. Hauser, S. Holroyd, K. Jagels, S. Leather, S. Moule, H. Norberczak, S. O'Neil, D. Ormond, C. Price, E. Rabbinowitsch, S. Rutter, M. Sanders, D. Saunders, K. Seeger, S. Sharp, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, L. Unwin, S. Whitehead, B. G. Barrell and D. J. Maskell: Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat Genet*, 35, 32-40 (2003)

100. Callanan, M. J., T. P. Beresford and R. P. Ross: Genetic diversity in the lactose operons of *Lactobacillus helveticus* strains and its relationship to the role of these strains as commercial starter cultures. *Appl Environ Microbiol*, 71, 1655-1658 (2005)

101. Callanan, M. J., R. P. Ross and T. P. Beresford: Insertion sequence elements as mediators of strain diversity in *Lactobacillus helveticus*. *Int J Food Microbiol*, 120, 120-123 (2007)

102. Jansen, R., J. D. van Embden, W. Gaastra and L. M. Schouls: Identification of a novel family of sequence repeats among prokaryotes. *Omics*, 6, 23-33 (2002)

103. Kamerbeek, J., L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal and J. van Embden: Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol*, 35, 907-914 (1997)

104. Hoe, N., K. Nakashima, D. Grigsby, X. Pan, S. J. Dou, S. Naidich, M. Garcia, E. Kahn, D. Bergmire-Sweat and J. M. Musser: Rapid molecular genetic subtyping of serotype M1 group A *Streptococcus* strains. *Emerg Infect Dis*, 5, 254-263 (1999)

105. Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero and P. Horvath: CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315, 1709-1712 (2007)

106. Schleifer, K. H. and W. Ludwig: Phylogeny of the genus *Lactobacillus* and related genera. *System. Appl. Microbiol.*, 18, 461-467 (1995)

107. Berger, B., R. D. Pridmore, C. Barretto, F. Delmas-Julien, K. Schreiber, F. Arigoni and H. Brussow: Similarity and differences in the *Lactobacillus acidophilus* group identified by polyphasic analysis and comparative genomics. *J Bacteriol*, 189, 1311-1321 (2007)

108. Nicolas, P., P. Bessieres, S. D. Ehrlich, E. Maguin and M. van de Guchte: Extensive horizontal transfer of core genome genes between two *Lactobacillus* species found in the gastrointestinal tract. *BMC Evol Biol*, 7, 141 (2007)

109. Boekhorst, J., R. J. Siezen, M. C. Zwahlen, D. Vilanova, R. D. Pridmore, A. Mercenier, M. Kleerebezem, W. M. de Vos, H. Brussow and F. Desiere: The complete genomes of *Lactobacillus plantarum* and *Lactobacillus johnsonii* reveal extensive differences in chromosome organization and gene content. *Microbiology*, 150, 3601-3611 (2004)

110. Canchaya, C., M. J. Claesson, G. F. Fitzgerald, D. van Sinderen and P. W. O'Toole: Diversity of the genus *Lactobacillus* revealed by comparative genomics of five species. *Microbiology*, 152, 3185-3196 (2006)

111. Bolotin, A., B. Quinquis, P. Renault, A. Sorokin, S. D. Ehrlich, S. Kulakauskas, A. Lapidus, E. Goltsman, M. Mazur, G. D. Pusch, M. Fonstein, R. Overbeek, N. Kyprides, B. Purnelle, D. Prozzi, K. Ngui, D. Masuy, F. Hancy, S. Burteau, M. Boutry, J. Delcour, A. Goffeau and P. Hols: Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol*, 22, 1554-1558 (2004)

112. Bringel, F. and J. C. Hubert: Extent of genetic lesions of the arginine and pyrimidine biosynthetic pathways in *Lactobacillus plantarum, L. paraplantarum, L. pentosus, and L. casei*: prevalence of CO(2)-dependent auxotrophs and characterization of deficient arg genes in *L. plantarum*. *Appl Environ Microbiol*, 69, 2674-2683 (2003)

113. Cai, H., B. T. Rodriguez, W. Zhang, J. R. Broadbent and J. L. Steele: Genotypic and phenotypic characterization of *Lactobacillus casei* strains isolated from different ecological niches suggests frequent recombination and niche specificity. *Microbiology*, 153, 2655-2665 (2007)

114. Molenaar, D., F. Bringel, F. H. Schuren, W. M. de Vos, R. J. Siezen and M. Kleerebezem: Exploring *Lactobacillus plantarum* genome diversity by using microarrays. *J Bacteriol*, 187, 6119-6127 (2005)

115. Liu, M., F. H. J. Enckevort and R. J. Siezen: Genome update: lactic acid bacteria genome sequencing is booming. *Microbiology*, 151, 3811-3814 (2005)

Key Words: *Lactobacillus*, Probiotic, Genome, Functional Genomics, Lactic Acid Bacteria, Review

Send correspondence to: Todd R. Klaenhammer, Dept. of Food, Bioprocessing and Nutrition Sciences, North Carolina State University, Raleigh, NC 27695, Tel: 919-515-2972, Fax: 919-513-0014, E-mail: trk@unity.ncsu.edu

http://www.bioscience.org/current/vol14.htm