Brazilian Journal of Microbiology (2011) 42: 147-157
ISSN 1517-8382

COMPARATIVE ANALYSIS OF TWO COMPONENT SIGNAL TRANSDUCTION SYSTEMS OF THE *LACTOBACILLUS ACIDOPHILUS* GROUP

Yanhua Cui[1]*; Xiaojun Qu[2]

[1]School of Food Science and Engineering, Harbin Institute of Technology Harbin 150090 P. R. China; [2]Institute of Applied Microbiology, Heilongjiang science academy, Harbin, 150010, P. R. China.

Submitted: April 22, 2009; Approved: August 23, 2010.

ABSTRACT

The *Lactobacillus acidophilus* group is a phylogenetically distinct group of closely related lactobacilli. Members of this group are considered to have probiotic properties and occupy different environmental niches. Bacteria generally sense and respond to environmental changes through two component systems (TCSs) which consist of a histidine protein kinase (HPK) and its cognate response regulator (RR). With the use of in silico techniques, the five completely sequenced *L. acidophilus* group genomes were scanned in order to predict TCSs. Five to nine putative TCSs encoding genes were detected in individual genomes of the *L. acidophilus* group. The *L. acidophilus* group HPKs and RRs were classified into subfamilies using the Grebe and Stock classification method. Putative TCSs were analyzed with respect to conserved domains to predict biological functions. Putative biological functions were predicted for the *L. acidophilus* group HPKs and RRs by comparing them with those of other microorganisms. Some of TCSs were putatively involved in a wide variety of functions which are related with probiotic ability, including tolerance to acid and bile, production of antimicrobial peptides, resistibility to the glycopeptide antibiotic vancomycin, and oxidative condition.

Key words: *Lactobacillus acidophilus* group, two component system, histidine protein kinase, response regulator protein, bioinformatics analysis

INTRODUCTION

The *Lactobacillus acidophilus* group ("acidophilus complex") is a phylogenetically distinct group of closely related lactobacilli, containing, among others, *Lactobacillus acidophilus*, *Lactobacillus johnsonii*, *Lactobacillus gasseri*, *Lactobacillus crispatus*, *Lactobacillus amylovorus*, *Lactobacillus gallinarum*, *Lactobacillus delbrueckii* subsp. *bulgaricus* (30). Several members of this group are considered to have probiotic properties. Strains of *L. acidophilus*, *L. johnsonii*, and *L. delbrueckii* subsp. *bulgaricus* have been extensively studied for their probiotic activities, pathogen inhibition, epithelial cell attachment, and immunomodulation. Members of these species can inhibit pathogen, prevent intestinal tract infections, improve the immune system, and reduce inflammatory or allergic reactions (5, 20, 29). At the same time, *L. acidophilus* and *L. delbrueckii* subsp. *bulgaricus* have the ability to alleviate lactose intolerance. So they have not only been widely used in the manufacture of fermented dairy but are also consumed as probiotic products (5, 20, 29).

***Corresponding Author. Mailing address:** School of Food Science and Engineering, Harbin Institute of Technology Harbin 150090 P. R. China.; E-mail: yhcui@hit.edu.cn

Survival during passage through the gastrointestinal (GI) tract of humans is generally considered a key feature for probiotics to preserve their expected health-promoting effects (5, 11). Survival of microorganisms during their transit through the GI tract requires the ability to sense and respond to the various and changing conditions present in the environment. Bacteria generally sense and respond to environmental changes through two component systems (TCSs). TCSs are present in the majority of Gram-positive and Gram-negative bacteria, and are one of the most important mechanisms for external environmental sensing and signal transduction (15, 37). TCSs are involved in controlling a wide variety of physiological processes, such as chemotaxis, biofilm formation, stress, osmolarity, quorum sensing, and virulence (15, 38). A typical TCS consists of a membrane-associated histidine protein kinase (HPK) and a cytoplasmic response regulator (RR). The former detects specific environmental signals and the latter regulates expression of genes.

Although the *L. acidophilus* group has received much attention in the past few years and some *L. acidophilus* group genomes have recently been sequenced and published (1, 19, 27, 35), only little research has been done on TCSs in this bacterial group. Only recently, LBA1524/LBA1525 in *L. acidophilus* was found to be related with acid tolerance, and LBA1430/LBA1431 with bile tolerance (3, 26). Since so little is known about TCSs in the *L. acidophilus* group, we scanned TCSs in five genomes of this group and predicted function of putative TCSs. At the same time, we compared the differences between five members of this group based on TCSs.

MATERIALS AND METHODS

Sequence information

Complete genome sequences of *L. acidophilus* NCFM, *L. gasseri* ATCC 33323, *L. johnsonii* NCC533, *L. delbrueckii* subsp. *bulgaricus* ATCC 11842, *L. delbrueckii* subsp. *bulgaricus* ATCC BAA365, and *Lactobacillus plantarum* WCFS1were obtained from the National Center for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov/ genomes).

Sequence analysis

Protein domain organizations were determined by SMART and Pfam (4, 31). TMHMM 2.0 was used to detect transmembrane helices (17). Multiple sequence alignments were created using the CLUSTAL X software (34). The evolutionary distances were calculated using the software package TREECON. The phylogenetic tree was generated by the neighbour-joining method using the software package TREECON (36).

Identification of HPKs and RRs

The genome sequences of the *L. acidophilus* group were searched for genes encoding putative HPKs and RRs by means of HMMER2.3.2 (http://hmmer.wustl.edu/). Consensus Protein Families Database (www.sanger.ac.uk/software/Pfam) sequences for HisKA (PFAM00512), HATPase_c (Pfam02518), and Response_reg (PFAM00072) were used in HMMER searches of each genome. The HisKA HMM and HATPase_c HMM were used to scan for the phosphoryl-accepting domain and highly conserved HATPase domain of HPKs, while the Response_reg HMM was used to scan for the highly conserved phosphoryl accepting domain of RRs. Putative functions were assigned to target genes manually by sequence comparison to an existing protein database using the NCBI protein–protein BLAST server (www.ncbi.nlm.nih.gov/ blast/blastp).

RESULTS AND DISCUSSION

The two component systems in the *L. acidophilus* group

The five completely sequenced genomes of the *L. acidophilus* group, including *L. acidophilus* NCFM, *L. gasseri* ATCC 33323, *L. johnsonii* NCC533, *L. delbrueckii* subsp. *bulgaricus* ATCC 11842 and ATCC BAA365 were scanned in order to predict TCSs by means of the Pfam HMMs HisKA, HATPase_c and Response_reg. The *L. plantarum* WCFS1 genome was predicted in the same way for comparative analysis. *L. plantarum* is a flexible and versatile species that is

encountered in a variety of environmental niches, including some dairy, meat, and many vegetable or plant fermentations. This flexible and adaptive behavior is reflected by the relatively large number of regulatory and transport functions, including 13 TCSs (16).

The five to nine putative HPKs containing a conserved histidine residue and a C-terminal HATPase domain, and five to nine putative RRs containing a RR receiver domain were detected in the genomes of the *L. acidophilus* group (Table 1). 14 HPKs and 15 RRs were predicted in *L. plantarum* WCFS1, which is significantly higher than those of the *L. acidophilus* group. The difference of the genomes' size is one of reasons. *L. plantarum* and the *L. acidophilus* group species, respectively, belong to facultative heterofermentatives, and obligately homofermentatives. *L. plantarum* has a comprehensive sugar metabolism. Its genome encodes all enzymes required for the glycolysis and phosphoketolase pathways. *L. plantarum* displays heterolactic fermentation or homolactic fermentation, depending on the environmental conditions. However, the members of *L. acidophilus* group have only glycolysis pathway and display homolactic fermentation. It is proposed that differences in their metabolism are linked to the number of TCSs. *L. plantarum* need more regulation and signaling systems in order to perform the different metabolism. At the same time, the results likely reflect adaptations of the *L. acidophilus* group to the stable and nutritionally rich milk environment or human gut, where fewer biosynthetic functions and less adaptive regulation are required.

Table 1. Number of HPK-RR pairs detected in the *Lactobacillus acidophilus* group genomes and *L. plantarum* genome

Strain	GenomeMb	CG%	HPK	RR	HK-RR	Orphans HPK	Orphans RR	GenBank number
Lactobacillus acidophilus NCFM	1.99	34.7	8	8	8	0	0	CP000033
Lactobacillus gasseri ATCC 33323	1.95	35.3	5	5	5	0	0	CP000413
Lactobacillus johnsonii NCC533	1.99	34.6	9	9	9	0	0	AE017198
L. delbrueckii subsp. *bulgaricus* ATCC 11842	1.86	49.7	6	6	5	1	1	CR954253
L. delbrueckii subsp. *bulgaricus* ATCC BAA-365	1.86	49.7	7	7	7	0	0	CP000412
Lactobacillus plantarum WCFS1	3.31	44.5	14	15	14	0	0	AL935263

Most of the TCSs are arranged in pairs, either with the HPK-encoding gene first followed by the RR-encoding gene, or vice versa. *L. delbrueckii* subsp. *bulgaricus* ATCC11842 contained a HPK and a RR gene not encoded in pairs. However, there is not single HPK or RR in the genome of ATCC BAA365. This difference was likely related to the origin of strains. The strain ATCC11842 was originally isolated from bulgarian yogurt by S. Orla Jensen in 1919, while the strain ATCC BAA365 was derived from a French starter culture CHCC757. Strains of different origin have adapted differently to fit specific niches during evolution. Similar results were obtained in thermophilic lactic acid bacteria *Streptococcus thermophilus* (28). Based on comparative genome hybridization analysis of 47 dairy *S. thermophilus* strains, the researchers revealed variable gene composition among *S. thermophilus* strains. It was presumed that there were frequent recombination or gene transfer within *S. thermophilus,* and some genes have disappeared and degenerated in the stable environment. At the same time, we find that the numbers of TCSs in *S. thermophilus* LMD-9 are smaller than those in strains CNRZ1066 and LMG18311.

Classification of HPKs and RRs

For all HPKs and RRs detected, the protein domain organization was analyzed using TMHMM, Pfam and SMART. The results of these analyses are shown in Fig 1 A

and 1B. All HPKs are classified in 7 groups according to protein domain organizations. Most of the HPKs are predicted to be membrane localized, consistent with the observation that localization of the sensor kinase to the membrane of the bacterial cells appears to be a general feature of most TCSs (38). Many of these HPKs contain previously described domains, including PAS domain, PAC domain and HAMP domains. PAS (Per-ARNT-Sim) domains monitor changes in redox potential, cellular oxygen, overall energy level of a cell, light, and small ligands. This domain is found in proteins regulating circadian rhythms and hypoxia responses as well as in input domains for TCSs. PAS domains are frequently followed by a 40-to 45-amino-acid PAC motif (33). PAS/PAC domains are widely distributed but are found primarily in proteins involved in signaling or regulation of transcription. The groups 6 and 7 contained PAS domain, including LGAS_0065, LJ0066, LBA0079, Ldb0136, Ldb0963, LBUL_0112 and LBUL_0873. HAMP domain is found in several ATP-binding proteins, for example histidine kinase, DNA gyrase B, topoisomerases, heat shock protein HSP90, phytochrome-like ATPases and DNA mismatch repair proteins (2). HAMP domains are found in the groups 3, 5 and 6. The precise functions of HAMP domains are unknown, however, mutation in the HAMP repeat region of *Neurospora crassa* is responsible for the most severe osmosensitivity and dicarboximide resistance phenotypes (21).

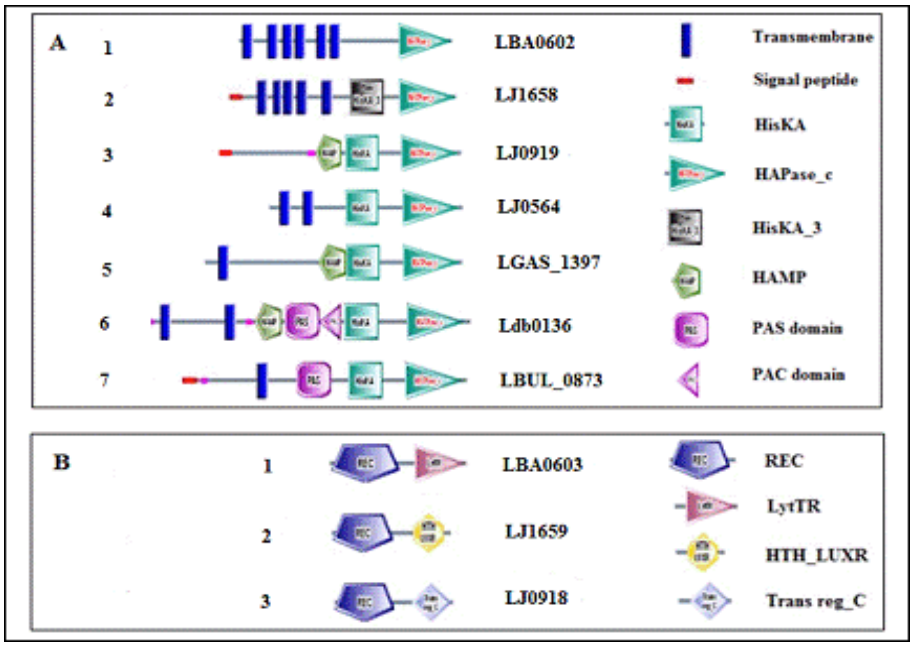

Figure 1. Scaled cartoon of domain structure of HPKs and RRs in *L. acidophilus* group

A Scaled cartoon of HPK domain structure for a representative protein from each group. 1 (LBA0602), 2 (LJ1658), 3 (LJ0919), 4 (LJ0564), 5(LGAS_1397), 6 (Ldb0136), 7 (LBUL_0873).

B Scaled cartoon of RR domain structure for a representative protein from each group. 1 (LBA0603), 2 (LJ1659), 3(LJ0918).

All RRs are classified in 3 groups according to protein domain organizations. Most output domains of RRs belonged to Trans_reg_C domain. The LJ1659 from *L. johnsonii* contains a typical LuxR-type HTH motif at the C terminus of proteins. This domain is a DNA-binding, helix-turn-helix (HTH) domain of about 65 amino acids, present in transcription regulators of the LuxR/FixJ family of response regulators. LuxR-type HTH domain proteins occur in a variety

of organisms. LuxR-type HTH regulators control a wide variety of activities in various biological processes, such as bioluminescence, virulence, spore formation, acetate metabolism (8-9, 13).

The LJ0766, LJ0918 of *L. johnsonii*, the LBA1798, LBA0603 of *L. acidophilus* and the LBUL_0021, Ldb0026 of *L. delbrueckii* subsp. *bulgaricus* contain LytTR type output domain. The LytTR domain is a DNA-binding, potential winged helix-turn-helix domain (~100 residues) present in a variety of bacterial transcriptional regulators of the *algR/agrA/lytR* family (23). LytTR domain is a type of DNA-binding domain and different from helix-turn-helix or winged-helix type output domain. This domain is distributed widely in low G+C Gram positive bacteria, and is involved in biosynthesis of extracellular polysaccharides, quorum sensing, bacteriocin peptide production (7, 18).

The putative HPKs and RRs were grouped in order to predict the subfamily of HPKs and RRs. Two bootstrapped NJ trees were constructed, an HPKs tree and an RRs tree (not shown). The HPKs tree was constructed with the *L. acidophilus* group HPKs phosphotransferase domains and highly conserved HATPase domain, while the RRs tree was constructed with all RR receiver domains.

In addition to this initial set of sequences, homologous sequences of other bacterial species were included to improve the resolution of both trees. Based on the two trees, the *L. acidophilus* group HPKs and RRs were classified into the subfamilies described by Grebe *et al.* (15). The results of classification are shown in Table 2. The HPKs typically contain two functionally and structurally distinct parts, a variable N-terminal sensor region and a conserved C-terminal kinase core domain. The latter have highly conserved residues called homology boxes, including the H-, N-, D-, F-, and G-boxes. The conserved boxes are presumed to play crucial roles in substrate binding, catalysis, and/or structure. Based on the presence and structure of the various homology boxes, Grebe and Stock made a comprehensive classification of HPKs (15). According to the criteria, the putative HPKs fell into five subfamilies (1a, 2a, 3a, 7, and 10). Most HPKs are members of the subfamily HPK_{1a}. This is the most common type HPK. The LJ1658 was unique and was found to belong to subfamily HPK_7. The HPK_7 subfamily has the following characteristics: the H-box is distinguished by the presence of a negatively charged group 2 residues upstream from the conserved histidine and a positively charged residue, usually an arginine, 8 residues upstream. The F-box is missing and the distance between the D- and the G-boxes is reduced. The LBA0602, LBA1799, LJ0448, LJ0764 and LBUL_0022 were classified into subfamily HPK_{10}. The HPK_{10} subfamily members commonly possess five to seven N-terminal transmembrane segments, and have no D-box. The subfamily 10 HPKs usually are related with quorum sensing.

Based on Grebe classification scheme, the putative RRs were grouped into 3 subfamilies (14-15). Most RRs were found to belong to OmpR subfamily. This subfamily appears to be the most abundant subfamily of RRs in the Gram positive bacteria whose genomes have been sequenced to date. The LBA0603, LBA1798, LJ0449, LJ0766, LBUL_0021, and Ldb0026 belonged to LytR subfamily, while the LJ1659 from *L. johnsonii* belonged to FixJ subfamily. Analysis of the two trees showed that the receiver domains of all RRs pairing to a HPK of a certain subfamily generally clustered together in the same branches of the RRs tree. For example, all RRs pairing with a subfamily HPK_{10} contained a HTH-DNA-binding domain of the LytTR family. These results were consistent with previous research that the HPK phosphotransferase domains, the cognate receiver domains and the RR output domains have evolved as integral units (15).

Function prediction of HPKs and RRs

To get functional annotation of the HPKs and RRs, bootstrapped NJ trees of HPKs and RRs were constructed with whole sequences of HPKs and RRs respectively (Fig 2-3). The phylogenetic analysis revealed nine major groups of the *L. acidophilus* group HPKs (Fig. 2). Five HPK groups (I, III, VII, VIII, IX) contain closely related sequences from all *L. acidophilus* group examined (Fig. 2). Thus, the sequences within each group are conserved in the *L. acidophilus* group

studied, may represent orthologs with common functions and likely involved in basic adaptation for environment. In contrast, the groups II, IV, V, VI contain the sequences from only some members. The results implied that the group II, IV, V, and VI were special for some members of the *L. acidophilus* group.

Lactobacillales-specific clusters of orthologous protein coding genes (LaCOGs) had been built using computational procedures in 12 sequenced *Lactobacillales* genomes (19). Most HPKs in group V belonged to LaCOG01758, including lp_0416, lp_3063, lp_1355, lp_3581, lp_3088, LJ0448, and LJ0764. The lp_0416 and lp_3581of *L. plantarum* were involved in QS, so we predicted the group V was related with QS. This result was consistent with previous research (10, 32). The HPKs in group IX belonged to LaCOG00289. The N-terminal domains of HPKs in group IX contained a HAMP domain and a PAS/PAC domain. The domain structure of group IX was similar to that of *Enterococcus faecalis* VicK (25). So these HPKs were likely related to resistance to glycopeptide antibiotics vancomycin.

The phylogenetic RRs tree revealed 9 major groups for the *L. acidophilus* group RRs (Fig. 3). Analysis of the two trees showed that all RRs pairing to a HPK of a certain group generally clustered together in the same branches of the RRs tree.

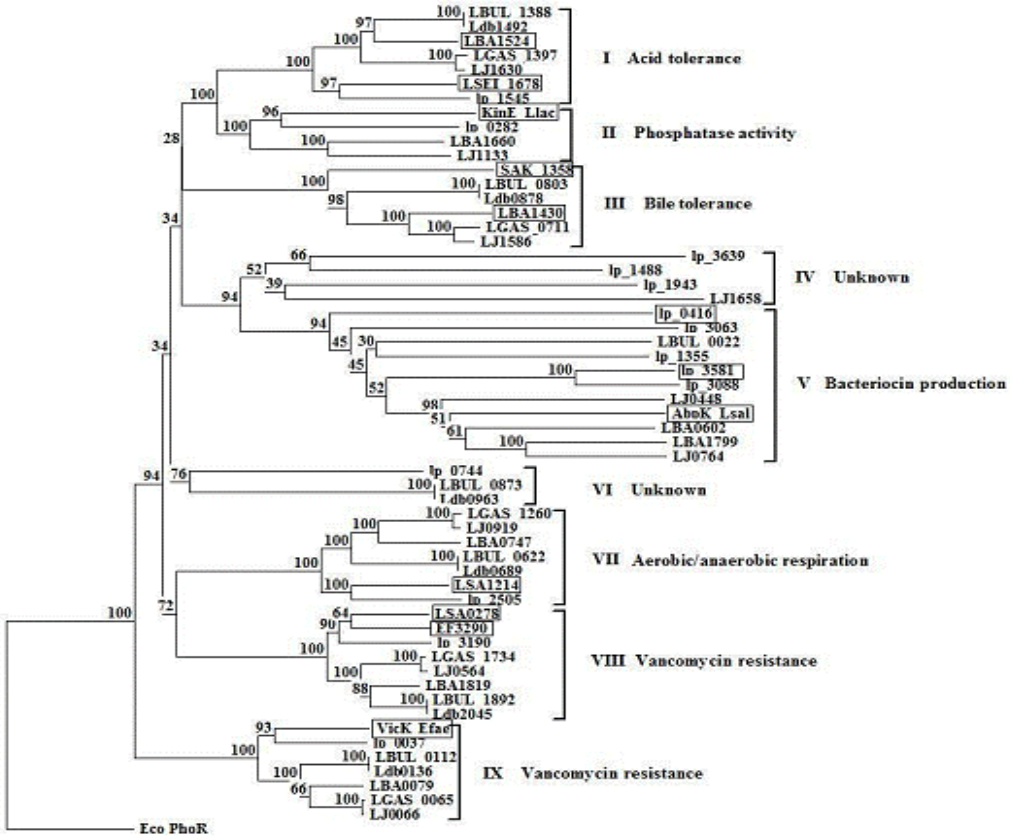

Figure 2. The phylogenetic tree of the *L. acidophilus* group HPKs.

HPK sequences were aligned by using ClustalW. The phylogenetic tree was generated by the neighbour-joining method using the software package TREECON. Eco_PhoR, HPK PhoR from *E. coli* (Protein code AP001050). The boxed HPKs indicate the functions of these sequences have been defined based on experiments (3, 10, 12, 22, 24-26). AbpK_Lsal (Protein code YP_536800), EF3290 (NP_816886), LSA0278 (YP_394892), VicK_Efae (AAO80993), KinE_Llac (YP_001032805), LSEI_1678 (ABJ70451), SAK_1358 (YP_329968), LSA1214 (YP_395826) are from *Lactobacillus salivarius* subsp. *salivarius*, *Enterococcus faecalis*, *Lactobacillus sakei* subsp. *sakei*, *E. faecalis, Lactococcus lactics* subsp *cremoris*, *Lactobacillus casei*, *Streptococcus agalactiae, L.sakei* subsp. *sakei*, respectively. The functions of HPKs are assigned based on sequences comparison with the functionally defined HPKs.

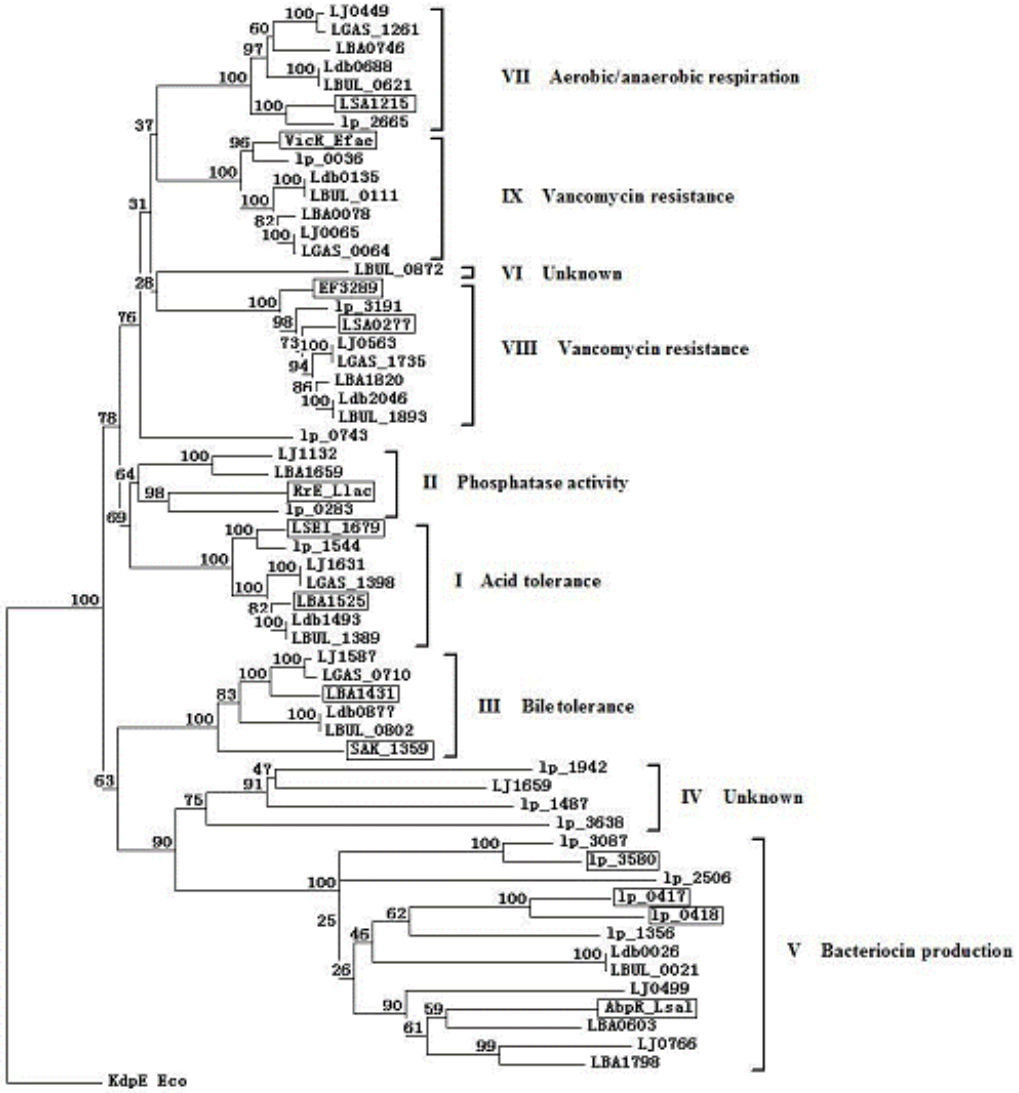

Figure 3. The phylogenetic tree of the *L. acidophilus* group RRs.

RR sequences were aligned by using ClustalW. The phylogenetic tree was generated by the neighbour-joining method using the software package TREECON. Eco_KdpE, RR KdpE from *E. coli* (Protein code YP851812). The boxed RRs indicate the functions of these sequences have been defined based on experiments (3, 10, 12, 22, 24-26). AbpR_Lsal (Protein code YP_536799), EF3289 (NP_816885), LSA0277(YP_394891), VicR_Efae (AAO80992), RrE_Llac (YP_001032807), LSEI_1679 (ABJ70452), SAK_1359 (YP_329969), LSA1215 (YP_395827) are from *Lactobacillus salivarius* subsp. *salivarius*, *Enterococcus faecalis*, *Lactobacillus sakei* subsp. *sakei*, *E. faecalis*, *Lactococcus lactics* subsp *cremoris*, *Lactobacillus casei*, *Streptococcus agalactiae*, *L.sakei* subsp. *sakei*, respectively. The functions of RRs are assigned based on sequences comparison with the functionally defined RRs.

To get a more specific functional annotation of the *L. acidophilus* group HPKs and RRs, they were compared with those of other bacterial species, using the NCBI BLAST server. Maintaining an E-value cut-off of 1e-44, we found a number of the *L. acidophilus* group TCSs to be similar to systems with a known biological function (Table 2, Suppl. Fig.1).

Table 2. The function prediction of the *Lactobacillus acidophilus* group HPKs and RRs

Strain	Locus	HPK class			Locus	RR class			HPK/RR	Homologous systems	Predicted function
		A	B	C		A	B	C	order		
L. acidophilus NCFM	LBA0079	6	IX	1a	LBA0078	3	IX	OmpR	RH	VicK/VicR 50%/77% Efae	Vancomycin resistance
	LBA0602	1	V	10	LBA0603	1	V	LytR	HR	AbpK /AbpR 33%/45% Lsal	Bacteriocin production
	LBA0747	3	VII	2a	LBA0746	3	VII	OmpR	RH	LSA1214/LSA1215 45%/ 63% Lsak	Aerobic/anaerobic
	LBA1430	4	III	3a	LBA1431	3	III	OmpR	HR	SAK_1358/ SAK_1359 38% / 57% Saga	Bile tolerance
	LBA1524	5	I	1a	LBA1525	3	I	OmpR	HR	LSEI_1678/ LSEI_1679 43%/73% Lcas	Acid tolerance
	LBA1660	4	II	1a	LBA1659	3	II	OmpR	RH	KinE/RrE 33%/ 47% Llac	Phosphatase activity
	LBA1799	1	V	10	LBA1798	1	V	LytR	RH	AbpK/AbpR 34%/38% Lsal	Bacteriocin production
	LBA1819	4	VIII	1a	LBA1820	3	VIII	OmpR	HR	LSA0278/ LSA0277 56%/85% Lsak	Vancomycin resistance
L. gasseri ATCC 33323	LGAS_0065	6	IX	1a	LGAS_0064	3	IX	OmpR	RH	VicK/VicR 51%/79% Efae	Vancomycin resistance
	LGAS_0711	4	III	3a	LGAS_0710	3	III	OmpR	RH	LBA1430/ LBA1431 64%/ 77% Laci	Bile tolerance
	LGAS_1260	3	VII	2a	LGAS_1261	3	VII	OmpR	HR	LSA1214/ LSA1215 46%/62% Lsak	Aerobic/anaerobic
	LGAS_1397	5	I	1a	LGAS_1398	3	I	OmpR	HR	LBA1524/LBA1525 55%/ 85% Laci	Acid tolerance
	LGAS_1734	4	VIII	1a	LGAS_1735	3	VIII	OmpR	HR	EF3290/ EF3289 51%/77% Efae	Vancomycin resistance
L. johnsonii NCC533	LJ0066	6	IX	1a	LJ0065	3	IX	OmpR	RH	VicK/VicR 51 %/80 % Efae	Vancomycin resistance
	LJ0448	1	V	10	LJ0449	1	V	LytR	HR	AbpK/ AbpR 34%/ 36% Lsal	Bacteriocin production
	LJ0564	4	VIII	1a	LJ0563	3	VIII	OmpR	RH	LSA0278/LSA0277 59%/83% Lsak	Vancomycin resistance
	LJ0764	1	V	10	LJ0766	1	V	LytR	HR	AbpK/AbpR 34%/38% Lsal	Lactacin F of
	LJ0919	3	VII	2a	LJ0918	3	VII	OmpR	RH	LSA1214/LSA1215 47%/62% Lsak	Aerobic/anaerobic
	LJ1133	4	II	1a	LJ1132	3	II	OmpR	RH	KinE/RrE 32%/46% Llac	Phosphatase activity
	LJ1586	4	III	3a	LJ1587	3	III	OmpR	HR	LBA1430/LBA1431 64%/ 77% Laci	Bile tolerance
	LJ1630	5	I	1a	LJ1631	3	I	OmpR	HR	LBA1524/LBA1525 56%/ 85% Laci	Acid tolerance
	LJ1658	2	IV	7	LJ1659	2	IV	FixJ	HR	SMU.1965c /SMU.1964c 40%/65% Smut	Unknown
L. delbrueckii subsp. *bulgaricus* ATCC 11842	Ldb0136	6	IX	1a	Ldb0135	3	IX	OmpR	RH	VicK/VicR 49%/75% Efae	Vancomycin resistance
	Ldb0689	3	VII	1a	Ldb0688	3	VII	OmpR	RH	LSA1214/ LSA1215 42%/66% Lsak	Aerobic/anaerobic
	Ldb0878	4	III	1a	Ldb0877	3	III	OmpR	RH	LBA1430/ LBA1431 44%/ 67% Laci	Bile tolerance
	Ldb1492	5	I	1a	Ldb1493	3	I	OmpR	HR	LBA1524/LBA1525 62%/ 91% Laci	Acid tolerance
	Ldb2045	4	VIII	1a	Ldb2046	3	VIII	OmpR	HR	EF3290/ EF3289 52%/80% Efae	Vancomycin resistance
	Ldb0963*	7	VI	1a						STRINF_01064 41% Sinf	Unknown
					Ldb0026#	1	V	LytR		NP_784991 28% Lplan	Bacteriocin production
L. delbrueckii subsp. *bulgaricus* BAA365	LBUL_0022	1	V	10	LBUL_0021	1	V	LytR	RH	NP_784990/NP_784991 33%/28% Lplan	Bacteriocin production
	LBUL_0112	6	IX	1a	LBUL_0111	3	IX	OmpR	RH	VicK/VicR 49%/75% Efae	Vancomycin resistance
	LBUL_0622	3	VII	1a	LBUL_0621	3	VII	OmpR	RH	LSA1214/LSA1215 42%/ 66% Lsak	Aerobic/anaerobic
	LBUL_0803	4	III	1a	LBUL_0802	3	III	OmpR	RH	LBA1430/ LBA1431 44%/ 67% Laci	Bile tolerance
	LBUL_0873	7	VI	1a	LBUL_0872	3	VI	OmpR	RH	STRINF_01064/CLOSCI_03935 41%/60%	Unknown
	LBUL_1388	5	I	1a	LBUL_1389	3	I	OmpR	HR	LBA1524/LBA1525 62%/ 91% Laci	Acid tolerance
	LBUL_1892	4	VIII	1a	LBUL_1893	3	VIII	OmpR	HR	LSA0278/ LSA0277 55%/82% Lsak	Vancomycin resistance

Lsak, *Lactobacillus sakei* subsp. *sakei* 23K; Lsal, *Lactobacillus salivarius* subsp. *salivarius* UCC118; Saga, *Streptococcus agalactiae* A909; Lcas, *Lactobacillus casei* BL23; Llac, *Lactococcus lactics* subsp *cremoris* MG1363; Efae, *Enterococcus faecalis* V583; Laci, *Lactobacillus acidophilus* NCFM; Lplan, *Lactobacillus plantarum* WCFS1; Lreu, *Lactobacillus reuteri* 100-23; Sinf, *Streptococcus infantarius* subsp. *infantarius* ATCC BAA-102; Csci, *Clostridium scindens* ATCC 35704; Smut, *Streptococcus mutans* UA159. *, Ldb0963 is orphan HPK. #, Ldb0026 is orphan RR.

A, according to protein domain organizations; B, according to phylogenetic analysis of HPKs or RRs; C, according to Grebe and Stock classification method.

Survival of bacteria is an important first step in the colonization of and probiotic contribution to the GI tract (5). The gastric acidity is the main obstacle to survival of bacteria. So the capacity of a microorganism to tolerate acidic pH is essential to the production and functionality of a probiotic culture. The TCS LBA1524/LBA1525 from *L. acidophilus* is identified to respond to acid (3). The insertional inactivation of the LBA1524 gene was found to reduce cell survival in pH 3.5. Thus the LBA1524/LBA1525 may aid the persistence and survival of microbes in an acidic environment. The TCSs LGAS_1397/LGAS_1398 and LJ1630/LJ1631 showed high identity to LBA1524/LBA1525. The HPKs and RRs of these TCSs showed 55%-56% and 85% identity to LBA1524 and LBA1525, respectively. *L. delbrueckii* subsp. *bulgaricus* is widely used as starter culture in the manufacture of yogurt and fermented milk products, so the strain need to tolerate the low pH in milk fermentation. The Ldb1492/Ldb1493 and LBUL_1388/LBUL_1389 in *L. delbrueckii* subsp. *bulgaricus* show high similarities to the LBA1524/LBA1525. The above mentioned TCSs were predicted to tolerate high acidic environment.

Because these *Lactobacillus* strains reside in the intestines, they must tolerate the presence of bile to survive in this environment. Bile is a multifaceted stressor, which can disrupt cell membranes and cause damage to DNA and proteins. The LBA1430/LBA1431 from *L. acidophilus* has been previously identified to respond to bile stress (26). We could also identify some TCSs putatively involved in tolerance bile. These TCSs are similar to LBA1430/LBA1431, including LGAS_0711/LGAS_0710, LJ1586/LJ1587, Ldb0878/Ldb0877 and LBUL_0803/LBUL_0802. The presence of these TCSs in the *L. acidophilus* group indicates an adaptation to the GI tract, enabling the bacteria to survive the acidic and bile-rich environments of the stomach and small intestine.

One of the properties of a probiotic strain is the ability to produce antimicrobial substances such as bacteriocins. *L. acidophilus* and *L. johnsonii* can produce a number of different bacteriocins, including lactacin F, lactacin, Acidophilucin A, and lactacin F. The genomes of *L. acidophilus* and *L. johnsonii* have revealed operons coding for bacteriocins. The bacteriocins may have an important role in inhibiting pathogenic bacteria of the human gut. We could also identify TCSs from *L. johnsonii* and *L. acidophilus* putatively involved in bacteriocin production and resistance. The LBA0602/LBA0603, LBA1799/LBA1798, LJ0448/LJ0449, and LJ0764/LJ0766 are similar to AbpK/AbpR which plays important role in the production of class II bacteriocins ABP-118 in *L. salivarius* (12). The LJ0448/ LJ0449 is unique and different from other putative TCSs relating to bacteriocin production. The GC content of the LJ0448 and LJ0449 genes are 22.4% and 24.3%. These GC contents are significantly lower than the average value of 34.6% observed for the entire *L. johnsonii* NCC533 genome. Interestingly, the LJ0448/ LJ0449 genes have not been found in other bacterial genomes, using Blastb. The results indicate that it may have been acquired recently via horizontal gene transfer.

We could also identify some TCSs from each member of the *L. acidophilus* group examined putatively involved in susceptibility to the glycopeptide antibiotic vancomycin (Table 2). These TCSs showed similarities to VicK/VicR in *Enterococcus faecalis* and HPK48/RRP48 in *L. sakei* (22, 25). Vancomycin is widely used to treat severe infections by Gram-positive bacteria. In lactic acid bacteria, the mechanism of resistance to vancomycin remains to be elucidated, although some lactic acid bacteria showed resistance to vancomycin.

Among putative TCSs, the TCS LJ1658/LJ1659 from *L. johnsonii* is unique. LJ1658 belonged to subfamily HPK_7, and LJ1659 contained a FixJ type output domain. Although a vast majority of QS-TCSs comprise HPK_{10} type HPK and LytR type RR, ComP/ComA of *Bacillus subtilis* is a QS-TCS that belong to HPK_7 and FixJ type RR (36). So we presumed LJ1658/LJ1659 was possibly involved in QS. The TCS was similar to a system of unknown function (SMU.1965c/ SMU.1964c) of *Streptococcus mutans* (6). Furthermore, the genes encoding the TCSs LJ1658/LJ1659 and SMU.1965c/ SMU.1964c appeared to share strong gene neighbourhood conservation. Based on the neighbouring genes, which encode putative (sugar) periplasmic transport systems, these TCSs

were putatively involved in host-microbe interactions, functioning in the utilization of nutrients in host.

By scanning five members of the *L. acidophilus* group genomes for TCSs, we have gained information about the capacity of these probiotic microorganisms to adapt to changes in their environment. The five to nine TCSs were predicted in the five genomes. These TCSs were involved in adapting to specific environment (e.g., acid tolerance, bile tolerance, aerobic/anaerobic respiration), resistance to glycopeptide antibiotics vancomycin, and production of bacteriocin. The results presented here provide a basis for future research on signal transduction mechanisms in the *L. acidophilus* group. At the same time, the results showed some TCSs were conserved in the *L. acidophilus* group, and other TCSs were specific for some lactic acid bacteria. The distribution of TCSs in the *L. acidophilus* group showed these lactic acid bacteria had adapted differently to fit their specific niches.

ACKNOWLEDGEMENTS

This work was supported by Natural Scientific Research Innovation Foundation in Harbin Institute of Technology (HIT. NSRIF. 2008.19), National Nature Science Foundation of China (Grant No.30901048), and Development Program for Outstanding Young Teachers in Harbin Institute of Technology (HITQNJS. 2007. 36).

REFERENCES

1. Altermann, E.; Russell, W.M.; Azcarate-Peril, M.A.; Barrangou, R.; Buck, B.L.; McAuliffe, O. (2005). Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc. Natl. Acad. Sci.U.S.A.*102 (11), 3906-3912.
2. Aravind, L.; Ponting, C.P. (1999). The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. *FEMS Microbiol. Lett.* 176, 111-116.
3. Azcarate-Peril, M.A.; McAuliffe, O.; Altermann, E.; Lick, S.; Russell, W.M.; Klaenhammer, T.R. (2005). Microarray analysis of a two-component regulatory system involved in acid resistance and proteolytic activity in *Lactobacillus acidophilus*. *Appl. Environ. Microbiol.* 71(10), 5794-5804.
4. Bateman, A.; Coin, L.; Durbin, R.; Finn, R.D.; Hollich,V.; Griffiths-Jones, S. (2004). The Pfam protein families database. *Nucleic Acids Res.* 32, D138-D141.
5. Bezkorovainy, A. (2001). Probiotics: determinants of survival and growth in the gut. *Am. J. Clin. Nutr.* 73, 399S-405S.
6. Biswas, I.; Drake, L.; Erkina, D.; Biswas, S. (2008). Involvement of sensor kinases in the stress tolerance response of *Streptococcus mutans*. *J Bacteriol*. 190(1), 68-77.
7. Carter, G.P.; Lyras, D.; Allen, D.L.; Mackin, K.E.; Howarth, P.M.; O'Connor, J.R.; Rood, J.I. (2007). Binary toxin production in *Clostridium difficil*e is regulated by CdtR, a LytTR family response regulator. *J. Bacteriol.* 189(20), 7290-7301.
8. Cramer, A.; Gerstmeir, R.; Schaffer, S.; Bott, M.; Eikmanns, B.J. (2006). Identification of RamA, a novel LuxR-Type transcriptional regulator of genes involved in acetate metabolism of *Corynebacterium glutamicum*. *J. Bacteriol.* 188(7), 2554-2567.
9. Crater, D.L.; Moran, C.P.Jr. (2001). Identification of a DNA binding region in GerE from *Bacillus subtilis*. *J. Bacteriol.* 183, 4183-4189.
10. Diep, D.B.; Johnsborg, O.; Risøen, P.A. (2001). Evidence for dual functionality of the operon *plnABCD* in the regulation of bacteriocin production in *Lactobacillus plantarum. Mol. Microbiol.* 41, 633-644.
11. Elli, M.; Callegari, M.L.; Ferrari, S.; Bessi, E.; Cattivelli, D.; Soldi, S.; Morelli, L.; Feuillerat, N.G.L.; Antoine, J.M. (2006). Survival of yogurt bacteria in the Human gut. *Appl. Environ. Microbiol.* 72(7), 5113-5117.
12. Flynn, S.; van Sinderen, D.; Thornton, G.M.; Holo, H.; Nes, I.F.; Collins, J. K. (2002). 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.
13. Fuqua, W.C.; Winans, S.C.; Greenberg, E.P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators . *J. Bacteriol.* 176, 269-275.
14. Galperin, M.Y. (2006). Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.* 188,4169-4182.
15. Grebe, T.W.; Stock, J.B. (1999). The histidine protein kinase superfamily. *Adv. Microb. Physiol.* 41, 139-227.
16. Kleerebezem, M.; Boekhorst, J.; van Kranenburg, R.; Molenaar, D.; Kuipers, O.P.; Leer, R. Tarchini, R.; Peters, S. A.; Sandbrink, H.M.; Fiers, M.W.E.J.; Stiekema,W.; Lankhorst, R.M.K.; Bron, P.A.; Hoffer, S.M.; Groot, M.N.N.; Kerkhoven, R.; de Vries, M.; Ursing, B.; de Vos, W.M.; Siezen, R.J. (2003). Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc. Natl. Acad. Sci. U.S.A.* 100(4), 1990-1995.
17. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305,567-580.
18. Lizewski, S.E.; Lundberg, D.S.; Schurr, M.J. (2002). The transcriptional regulator AlgR is essential for *Pseudomonas aeruginosa pathogenesis*.

*Infect Immun.*70, 6083-6093.

19. Makarova, K.; Slesarev, A.; Wolfa, Y.; Sorokina, A.; Mirkinc, B.; Koonina, E. (2006). Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 103(42), 15611-15616.
20. Mercenier, A.;Pavan, S.; Pot, B. (2002) . Probiotics as biotherapeutic agents: present knowledge and future prospects. *Curr. Pharm. Des.* 8, 99-110.
21. Miller, T.K.; Renault, S.; Selitrennikoff, C.P. (2002). Molecular dissection of alleles of the osmotic-1 locus of *Neurospora crassa*. *Fungal Genet. Biol.* 35,147-155.
22. Morel-Deville, F.; Fauvel, F.; Morel, P. (1998). Two-component signal-transducing systems involved in stress responses and vancomycin susceptibility in *Lactobacillus sakei*. *Microbiology*.144, 2873-2883.
23. Nikolskaya, A.N.; Galperin, M.Y. (2002). A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic Acids Res.* 30, 2453-2459.
24. O'Connell-Motherway, M.; van Sinderen, D.; Morel-Deville, F.; Fitzgerald, G.F.; Ehrlich, S. (2000). Six putative two-component regulatory systems isolated from *Lactococcus lactis* subsp. *cremoris* MG1363. *Microbiology* 146, 935-947.
25. Paulsen, I.T.; Banerjei, L.; Myers, G.S.A. (2003). Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science.* 299 (5615), 2071-2074.
26. Pfeiler, E.A.; Azcarate-Peril, M.A.; Klaenhammer, T.R. (2007). Characterization of a novel bile-Inducible operon encoding a two-component regulatory system in *Lactobacillus acidophilus*. *J. Bacteriol.* 189(13), 4624-4634.
27. Pridmore, R.D.; Berger, B.; Desiere, F.; Vilanova, D.; Barretto, C.; Pittet, A.C. (2004). The genome sequence of the probiotic intestinal bacterium *Lactobacillus johnsonii* NCC 533. *Proc. Natl. Acad. Sci. U.S.A.* 101(8), 2512-2517.
28. Rasmussen, T.B.; Danielsen, M.; Valina, O.; Garrigues, C.; Johansen, E.; Pedersen, M.B. (2008). *Streptococcus thermophilus* core genome: comparative genome hybridization study of 47 strains. *Appl. Environ. Microbiol.* 74(15), 4703-4710.
29. Saarela, M.; Mogensen, G.; Fondén, R.; Mättö, J.; Mattila-Sandholm, T. (2000) Probiotic bacteria: safety, functional and technological properties. *J Biotechnol.* 84:197-215.
30. Schleifer, K.H.; Ludwig, W. (1995). Phylogeny of the genus *Lactobacillus* and related genera. *Syst. Appl. Microbiol.* 18, 461-467.
31. Schultz, J.; Milpetz, F.; Bork, P.; Ponting, C.P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5857-5864.
32. Sturme, M.H.J.; Nakayama, J.; Molenaar, D.; Murakami, Y.; Kunugi, R.; Fujii, T.;Vaughan, E.E.; Kleerebezem, M.; de Vos, W.M. (2005). 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(15), 5224-5235.
33. Taylor, B.L.; Zhulin, I.B. (1999). PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* 63, 479-506.
34. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. (1997). The clustalx windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876-4882.
35. van de Guchte, M.; Penaud, S.; Grimaldi, C.; Barbe, V.; Bryson, K.; Nicolas, P.; Robert, C.; Oztas, S.; Mangenot,S.; Couloux, A.; Loux, V.; Dervyn, R.; Bossy, R.; Bolotin, A.; Batto, J.M.; Walunas, T.; Gibrat, J.F.; Bessières, P.; Weissenbach, J.; Ehrlich, S.D. ; Maguin, E. (2006). The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9274-9279.
36. van der Peer, Y.; de Wachter, R. (1994). TREECON for windows: a software package for construction and drawing of evolutionary trees for the microsoft windows environment. *Comput. Appl. Biosci.* 10, 569-570.
37. Weinrauch, Y.; Penchev, R.; Dubnau, E.; Smith, I.; Dubnau, D. (1990). A *Bacillus subtilis* regulatory gene product for genetic competence and sporulation resembles sensor protein members of the bacterial two-component signal-transduction systems. *Genes Dev.* 4, 860-872.
38. West, A.H.; Stock, A.M. (2001). Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem. Sci.* 26, 369-376.