



**HAL**  
open science

# Identification des gènes impliqués lors de l'établissement de *Lactobacillus casei* dans l'intestin et caractérisation de l'opéron LSEI\_0219-0221

Hélène Scornec

## ► To cite this version:

Hélène Scornec. Identification des gènes impliqués lors de l'établissement de *Lactobacillus casei* dans l'intestin et caractérisation de l'opéron LSEI\_0219-0221. Biochimie, Biologie Moléculaire. Université de Bourgogne, 2014. Français. NNT : 2014DIJOS088 . tel-03018674

**HAL Id: tel-03018674**

**<https://theses.hal.science/tel-03018674>**

Submitted on 23 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Université de Bourgogne, AgroSup Dijon**

UMR A 02102 Procédés Alimentaires et Microbiologiques

Equipe Procédés Microbiologiques et Biotechnologiques

Ecole doctorale ES – Environnements – Santé

## **THESE DE DOCTORAT**

Discipline : Sciences de l'Alimentation

Spécialité : Microbiologie, Biologie Moléculaire

Présentée par

**Hélène SCORNEC**

Pour obtenir le titre de Docteur de l'Université de Bourgogne

### **Identification des gènes impliqués lors de l'établissement de *Lactobacillus casei* dans l'intestin et caractérisation de l'opéron LSEI\_0219-0221**

**Soutenue publiquement le 4 novembre 2014 devant la commission d'examen suivante :**

Mme Véronique MONNET, DR	INRA, Jouy-en-Josas	Rapporteur
M. Jacques RAVEL, Pr	University of Maryland	Rapporteur
M. Dominique GARMYN, MCF	INRA, UB, Dijon	Examineur
Mme Nathalie LEBLOND-BOURGET, Pr	Université de Lorraine	Examineur
M. Philippe J. SANSONETTI, Pr	Institut Pasteur, Paris	Examineur
M. Jean-François CAVIN, Pr	AgroSup Dijon, UB	Directeur de thèse
Mme Hélène LICANDRO-SERAUT, Dr	AgroSup Dijon, UB	Co-encadrante de thèse

## Remerciements

Je souhaite tout d'abord remercier Véronique Monnet et Jacques Ravel de m'avoir fait l'honneur d'accepter de juger ce travail de thèse.

Je tiens aussi à remercier M. Patrick Gervais et M. Jean-Marie-Perrier Cornet de m'avoir accueillie au sein de leur laboratoire.

Je remercie tout particulièrement toutes les personnes qui ont participé à l'encadrement de cette thèse : Jean-François Cavin, Hélène Licandro-Seraut, Philippe Sansonetti et Thierry Pédrón.

Je remercie Jean-François, mon directeur de thèse, pour ses nombreux conseils scientifiques et techniques et pour la liberté et l'autonomie qu'il m'a laissé avoir au cours de mes recherches. Merci aussi pour les pauses café agrémentées de discussion et de croissants lors de ma rédaction de thèse.

Je tiens à remercier Hélène, pour son encadrement, ses précieux conseils, sa patience, sa confiance et son enthousiasme. Je la remercie de m'avoir fait partager son expérience et son savoir-faire et de m'avoir encouragée dans les moments de doute. Merci pour la relecture et pour les corrections de ce mémoire. J'ai pu apprécier les nombreuses discussions scientifiques et personnelles que nous avons eues tout au long de cette thèse.

J'adresse mes remerciements les plus sincères à Philippe pour son implication et l'intérêt qu'il a manifesté dans l'encadrement de ma thèse, ainsi que pour ses conseils d'orientation de recherche. Merci aussi d'avoir initié et permis les collaborations avec les différentes équipes de l'Institut Pasteur, la chance de travailler avec plusieurs équipes de recherche s'est avérée être une expérience très enrichissante. Et surtout, merci de me faire l'honneur d'être examinateur de mon travail de thèse.

Je remercie Thierry pour son accueil chaleureux et sa gentillesse lors de mes séjours à l'Institut Pasteur, pour les manips de microbiologie qu'il a bien voulu faire même s'il dit toujours qu'il n'est pas microbiologiste. Je le remercie aussi pour la relecture des différents rapports et articles ainsi que pour ses conseils avisés.

Je tiens aussi à remercier Nathalie Leblond-Bourget et Dominique Garmyn d'avoir accepté d'être examinateur de ma thèse et pour leurs précieux conseils et encouragements lors des différents comités de suivi.

Je remercie également Christiane Bouchier et Magali Tichit pour leur accueil à la plateforme génomique et leur formation sur la technique de séquençage.

Je tiens tout particulièrement à remercier tous les membres du laboratoire de microbiologie qui par leur bonne humeur et leur gentillesse ont rythmé mes trois années de thèse.

Merci à Yves Waché, maître Yoda à sa façon, pour sa sagesse, sa diplomatie et pour ses conseils scientifiques et personnels.

Merci à Florence Husson pour sa bonne humeur, son rire communicatif et pour avoir fait de du deuxième étage un endroit riche d'humour et de sympathie.

Merci à Christine de nous avoir accueilli tous les midis dans son bureau pour notre pause déjeuner et pour la gestion des différents repas de labo.

Merci aux thésards et post-doc du bureau d'à côté : Hanh, Thu, Manhal, Cynthia et Jean-Luc pour leur partage d'expérience et pour nos fous rires.

Merci aux thésards et Post-doc d'autres environs : Célia, Jenna, Charline, Cyril, Maud et Magali pour nos différentes conversations professionnelles et personnelles qui se sont révélées motivantes et encourageantes.

Merci à ceux qui ont partagé le bureau 217 depuis le début de thèse : Giorgia, merci de m'avoir fait redécouvrir ton pays et merci pour les différents moments que nous avons pu passer ensemble. Chi, partie avant mon arrivée au labo et revenue quelques mois avant mon départ, ça a été une chance de faire ta connaissance et de converser avec toi, merci aussi de nous avoir quasiment nourris tout l'été avec les délicieuses figes de ton jardin. Et enfin merci à ceux qui ont partagé ce bureau depuis le début de la thèse, Dat et Bao Ngoc, merci pour nos discussions parfois existentielles, d'avoir accompagné mes longues journées d'été, et d'avoir partagé vos souvenirs du Vietnam, j'espère qu'on aura l'occasion d'y aller un jour ensemble.

Merci aux différents stagiaires avec qui j'ai eu la chance de travailler.

Merci à mes amis pour leur soutien inconditionnel et leurs encouragements, j'ai de la chance de vous avoir.

Merci à ma famille, mon frère et mes parents, qui sans faille ont soutenu tout mon parcours, et plus particulièrement un grand merci à ma maman, ayant le rôle à la fois de mère et de confidente, d'avoir fait de moi la personne que je suis aujourd'hui.

Merci à Charles (mon loulou) de m'avoir soutenue, épaulée, encouragée et surtout pour ta patience quotidienne, je te dois beaucoup.

# Table des matières

AVANT-PROPOS	1
SYNTHESE BIBLIOGRAPHIQUE	4
<b>I. Commensales –Ecologie intestinale</b>	<b>5</b>
I.1. Le tractus gastro-intestinal	5
I.2. Composition de la flore du tractus gastro-intestinal	5
I.3. Fonctions/Rôle de la flore du tractus gastro-intestinal	7
<b>II. Mécanismes d’adaptation par la régulation transcriptionnelle</b>	<b>8</b>
II.1. Les mécanismes de la transcription	8
II.2. La régulation transcriptionnelle	9
II.2.1. Les systèmes à un composant	10
II.2.2. Les systèmes à deux composants	10
II.2.3. Les systèmes à multiples composants	11
II.3. Les protéines d’un système à deux composants	13
II.3.1. Le composant Histidine Kinase	13
Classification des HKs	13
Composition des HKs	16
II.3.2. Le composant Response Regulator	17
Le domaine receveur	17
Le domaine effecteur	18
RR avec un domaine effecteur de liaison à l’ADN	18
RR avec un domaine effecteur de liaison à l’ARN	21
RR avec un domaine effecteur de liaison à une protéine	22
RR avec un domaine effecteur enzymatique	22
II.3.3. Transmission du signal de l’HK vers son RR et activité	22
Mécanismes de régulation	23
Activateurs transcriptionnels	23
Répresseurs transcriptionnels	23
Cross-talk entre RRs et HKs	24
II.4. Les systèmes à deux composants de <i>L. casei</i> ATCC 334	25
II.4.1. Détection des températures basses	28
II.4.2. Métabolisme de la paroi	28
II.4.3. Carence en phosphate	28
II.4.4. Résistance au stress et compétence	29
II.4.5. Quorum sensing-bactériocine	29
II.4.6. Utilisation de l’acide L-Malique	30
II.4.7. Résistance aux peptides antimicrobiens	30
II.4.8. Résistance au stress acide	31

<b>III. Architecture et rôle de la surface cellulaire des Lactobacilles dans les interactions hôte-microbe</b>	<b>31</b>
III.1. Le peptidoglycane	33
III.1.1. Structure et composition du peptidoglycane	33
III.1.2. Les Penicillin-Binding Proteins (PBPs)	35
III.1.4. Les PBPs de <i>Lactobacillus casei</i> ATCC 334	39
III.2. Les acides téichoïques	40
III.3. Les polysaccharides de la paroi	43
III.4. Les pili et flagelles	43
III.5. Les protéines de surface	45
III.5.1. Les protéines ancrées dans la membrane cytoplasmique	45
III.5.2. Les lipoprotéines	47
III.5.3. Les protéines ancrées à la paroi avec un domaine LPXTG.	48
III.5.4. Les protéines liées de manière non-covalente à la paroi	49
Les protéines à domaine WxL	49
Les protéines à domaine LysM	50
Les protéines à domaine SH3	50
Les protéines à domaine putatif de liaison au peptidoglycane	50
Les protéines de la couche S	50
Les protéines à domaine de liaison à la choline	51
<b>CONTEXTE, OBJECTIFS ET STRATEGIES</b>	<b>52</b>
<b>MATERIELS ET METHODES</b>	<b>55</b>
<b>I. Souches bactériennes et conditions de culture</b>	<b>56</b>
<b>II. Techniques générales de manipulations des acides nucléiques</b>	<b>56</b>
II.1. Extraction des acides nucléiques	56
II.1.1. Extraction d'ADN total	56
II.1.2. Extraction d'ADN plasmidique	57
II.1.3. Extraction des ARN totaux	57
II.2. Amplification de l'ADN par PCR	57
II.3. RT-qPCR (Reverse Transcription quantitative Polymerase Chain Reaction)	58
II.4. PCR quantitative	60
II.5. Séquençage d'ADN	60
<b>III. Techniques de transformation</b>	<b>60</b>
III.1. Transformation de <i>L. casei</i>	60
III.2. Transformation d' <i>E. coli</i>	61
<b>IV. Construction plasmidique</b>	<b>61</b>
<b>V. Techniques de manipulation des protéines</b>	<b>62</b>
V.1. Obtention d'extrait brut acellulaire	62
V.1.1. Préparation des cellules	62
V.1.2. Casse des cellules au disrupteur	62
V.1.3. Casse des cellules aux billes de verre	63
V.1.4. Extraction des protéines de surface liées de manière non covalente	63

V.2. Gels protéiques	63
V.2.1.SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis)	63
V.2.2. BN-PAGE (Blue Native PAGE)	64
V.3. Surexpression, purification et phosphorylation de protéines	64
V.3.1. Surexpression de protéines	64
V.3.2. Purification de protéines	65
V.3.3. Clivage du His-tag	65
V.3.4. Autophosphorylation et transfert de phosphate	65
<b>VI. EMSA (Electrophoretic Mobility Shift Assay)</b>	<b>66</b>
<b>VII. Autolyse induite au Triton X-100</b>	<b>67</b>
<b>VIII. Mesure de la variation de la fluidité membranaire</b>	<b>68</b>
<b>IX. Microscopie électronique à transmission (TEM)</b>	<b>70</b>
<b>X. Tests d'adhésion</b>	<b>70</b>
X.1. Adhésion microbienne aux solvants (MATS, Microbial Adhesion To Solvent)	70
X.2. Adhésion à la mucine	71
X.3. Adhésion à une surface abiotique	71
<b>XI. Compétition souche parentale et mutants</b>	<b>72</b>
XI.1. Compétition <i>in vivo</i> dans les anses iléales de lapin	72
XI.2. Compétition <i>in vitro</i> dans différents milieux	72
<b>XII. Sensibilité aux antibiotiques</b>	<b>73</b>
<b>XIII. Analyse et traitement des données bioinformatiques</b>	<b>73</b>
<b>RESULTATS</b>	<b>74</b>
<b>I. Mutagenèse par transposition de <i>L. casei</i> et identification des sites d'insertion du transposon</b>	<b>76</b>
I.1. Mutagenèse par transposition de <i>L. casei</i>	76
I.2. Identification des sites d'insertion du transposon des mutants de <i>L. casei</i>	85
<b>II. Identification des gènes impliqués lors de l'établissement de <i>L. casei</i> dans l'intestin.</b>	<b>92</b>
<b>III. Caractérisation fonctionnelle de l'opéron LSEI_0219-0221</b>	<b>114</b>
<b>CONCLUSION GENERALE</b>	<b>149</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b>	<b>158</b>
<b>ANNEXES</b>	<b>169</b>

# Table des illustrations

## Figures

Figure 1 : Colonisation microbienne du tractus gastro-intestinal humain.....	5
Figure 2 : Représentation schématique du complexe d'initiation de la transcription. ....	9
Figure 3 : Représentation schématique d'un système à deux composants. ....	11
Figure 4 : Système de phosphorelais impliqué dans la sporulation de <i>B. subtilis</i> . ....	12
Figure 5 : Model de transduction du signal du système phosphorelay RcsCDB. ....	13
Figure 6 : Représentation des différents domaines constituant une histidine kinase.....	15
Figure 7 : Représentation schématique des trois groupes de HK. ....	16
Figure 8 : Représentation des différents domaines constituant un response regulator .....	17
Figure 9 : Organisation spatiale schématique du domaine REC des RR. ....	18
Figure 10 : Représentation schématique des motifs HTH et wHTH de liaison à l'ADN. ....	19
Figure 11 : Représentation tridimensionnelle du domaine effecteur d'OmpR d' <i>E. coli</i> .....	19
Figure 12 : Représentation tridimensionnelle du complexe entre l'ADN et le domaine effecteur LytTR d'AgrA de <i>S. aureus</i> . ....	20
Figure 13 : Représentation de la structure tridimensionnelle du complexe entre l'ADN et le domaine AraC-HTH de MarA d' <i>E. coli</i> . ....	21
Figure 14 : Représentation de la structure tridimensionnelle de la protéine AmiR.....	21
Figure 15 : Schématisation des différents mécanismes d'activation et de répression de l'initiation de la transcription par les régulateurs transcriptionnels. ....	24
Figure 16 : Communication entre les différentes protéines des TCS. ....	25
Figure 17 : Représentation de la position des TCSs sur le génome de <i>L. casei</i> ATCC 334.....	26
Figure 18 : Représentation des voies de signalisation senseur et de résistance de <i>L. casei</i> aux PAMs. ....	31
Figure 19 : Représentation schématique de l'enveloppe cellulaire des Lactobacilles .....	32
Figure 20 : Variation de la composition du pentapeptide du peptidoglycane.....	34
Figure 21 : Exemple de pont peptidique. ....	34
Figure 22 : Représentation de la structure du peptidoglycane de <i>L. casei</i> BL23.....	35
Figure 23 : Représentation des différentes fonctions des PBPs dans la biosynthèse du peptidoglycane chez <i>E. coli</i> . ....	36
Figure 24 : Représentation de la composition des WTA et des LTA de type I. ....	42
Figure 25 : Modèle d'incorporation des D-Alanyl esters sur les LTA.....	42
Figure 26 : Représentation du mécanisme d'assemblage des pili hétérotrimériques.....	44
Figure 27 : Représentation du processus de translocation de protéines.....	46
Figure 28 : Représentation schématique de la synthèse des lipoprotéines. ....	47

Figure 29 : Ancrage des protéines à motif LPxTG.....	49
Figure 30 : Logo représentant les variations du motif LPxTG chez <i>L. casei</i> ATCC 334.....	49
Figure 31 : Schéma des principaux objectifs du travail de thèse et des stratégies mises en œuvre.....	54
Figure 32 : Localisation des différentes sondes utilisées pour réaliser les EMSAs.....	67
Figure 33 : Sondes fluorescentes permettant d'évaluer la fluidité membranaire.....	69
Figure 34 : Adhésion des mutants de l'opéron ( $\Delta 0219$ , $\Delta 0220$ et $\Delta 0221$ ) et de <i>L. casei</i> ATCC 334 sur plaques recouvertes ou non de mucine. ....	143
Figure 35 : Phosphorylation du RR0219 en présence d'acétyl-phosphate.....	144
Figure 36 : Représentation schématique du gène LSEI_0219, de la protéine RR0219 et des différents domaines la constituant.....	144
Figure 37 : EMSA de l'extrait protéique cytoplasmique de <i>L. casei</i> ATCC 334 avec trois sondes ADN représentant la région intergénique des gènes LSEI_0218 et LSEI_0219.....	145
Figure 38 : Représentation des motifs putatifs de liaison à l'ADN du RR0219.....	145
Figure 39 : Taux de transcrits du gène LSEI_0219 obtenu après 1h dans les différentes conditions testées. ....	147
Figure 40 : Fluidité membranaire de <i>L. casei</i> en différentes conditions par mesure d'anisotropie de fluorescence.....	147
Figure 41 : Rôles démontrés et probables de l'opéron LSEI_0219-0221 dans la surface cellulaire de <i>L. casei</i> .....	157

## Tableaux

Tableau 1 : Représentation des différentes boxes retrouvées dans le domaine transmetteur des histidine kinases. ....	15
Tableau 2 : Récapitulatif des caractéristiques des TCS de <i>L. casei</i> ATCC 334.....	27
Tableau 3 : Classification des PBPs de sept bactéries.....	36
Tableau 4 : Fonctions de différentes PBPs dans différentes souches.....	38
Tableau 5 : Comparaison du nombre de chaque COG représentant les PBPs chez différentes souches bactériennes. ....	40
Tableau 6 : Concentration des antibiotiques utilisés.....	56
Tableau 7 : Amorces utilisées pour la construction des plasmides de surexpression des systèmes à deux composants.....	58
Tableau 8 : Souches microbiennes, vecteurs et plasmides utilisés dans ce travail.....	62
Tableau 9 : Amorces utilisées pour synthétiser les différentes sondes. ....	67

## Liste des abréviations

<b>ABC</b> : ATP-Binding Casette	<b>LTA</b> : lipoteichoic acids (acides lipotéichoïques)
<b>ADN</b> : acide désoxyribonucléique	<b>LTTR</b> : régulateur de transcription de type LysR
<b>ADNc</b> : ADN complémentaire	<b>MATS</b> : microbial adhesion to solvent
<b>ADP</b> : Adénosine Di-Phosphate	<b>MRS</b> : Man, Rogosa, Sharpe
<b>AgS</b> : Antigène de Surface	<b>NAG</b> : acide <i>N</i> -acétyl-glucosamine
<b>AMP</b> : Antimicrobial peptide (peptide antimicrobien)	<b>NAM</b> : acide <i>N</i> -acétyl-muramique
<b>ARN</b> : acide ribonucléique	<b>OMS</b> : Organisation Mondiale de la Santé
<b>ARNm</b> : ARN messenger	<b>PAGE</b> : PolyAcrylamide Gel Electrophoresis
<b>Asp</b> : acide aspartique	<b>PAI</b> : peptide auto-inducteur
<b>ASS</b> : ADN de sperme de saumon	<b>pb</b> : paire de bases
<b>ATCC</b> : American Type Culture Collection	<b>PBP</b> : Penicillin-Binding Protein
<b>ATP</b> : Adénosine Tri-Phosphate	<b>PBS</b> : Phosphate Buffer Saline
<b>BN</b> : Blue Native	<b>PCR</b> : Polymerase Chain Reaction
<b>BSA</b> : Bovin Serum Albumin (sérum albumine bovine)	<b>PG</b> : peptidoglycane
<b>CIP</b> : Collection de l'Institut Pasteur	<b>PGH</b> : peptidoglycan hydrolase (hydrolase du peptidoglycane)
<b>COG</b> : Cluster of Orthologous Groups	<b>RNAP</b> : RNA polymerase (ARN polymérase)
<b>CPS</b> : capsular polysaccharide (polysaccharide de la capsule)	<b>RNase</b> : ribonucléase
<b>DAP</b> : diaminopimelic acid (acide diaminopimélique)	<b>RR</b> : Response Regulator
<b>DEPC</b> : diéthyl pyrocarbonate	<b>RT</b> : Reverse Transcription
<b>dNTP</b> : désoxyribonucléotide	<b>SDS</b> : Sodium Dodecyl Sulfate
<b>DO</b> : Densité Optique	<b>STM</b> : Signature Tagged Mutagenesis
<b>DTT</b> : dithiothreitol	<b>TA</b> : Teichoic Acid (acide téichoïque)
<b>EMSA</b> : Electrophoretic Mobility Shift Assay	<b>TCS</b> : Two-Component System (système à deux composants)
<b>EPS</b> : exopolysaccharides	<b>TEM</b> : Transmission Electron Microscopy
<b>His</b> : histidine	<b>TGI</b> : Tractus Gastro-Intestinal
<b>HK</b> : Histidine Kinase	<b>TLR</b> : Toll-Like Receptor
<b>HLH</b> : Helix-Loop-Helix (hélice-boucle-hélice)	<b>TMR</b> : transmembranaire
<b>HMW</b> : High Molecular Weight (haut poids moléculaire)	<b>Tris</b> : Tris-(hydrométhyl)aminométhane
<b>HPt</b> : histidine phosphotransférase	<b>UFC</b> : Unité Formant Colonies
<b>HTH</b> : Helix-Turn-Helix (hélice-tour-hélice)	<b>wHTH</b> : winged Helix-Turn-Helix (hélice-tour-hélice ailée)
<b>kDa</b> : kiloDalton	<b>WPS</b> : wall polysaccharide (polysaccharide de la paroi)
<b>KEGG</b> : Kyoto Encyclopedia of Genes and Genomes	<b>WTA</b> : Wall Teichoic Acids (acides téichoïques de la paroi)
<b>LAB</b> : Lactic Acid Bacteria (bactéries lactiques)	
<b>LB</b> : Luria-Bertani	
<b>LMW</b> : Low Molecular Weight (faible poids moléculaire)	

## Liste des travaux

### Articles

**Scornec H**, Tichit M, Bouchier C, Pédrón T, Cavin J-F, Sansonetti PJ, Licandro-Seraut H. **2014**. Rapid 96-well plates DNA extraction and sequencing procedures to identify genome-wide transposon insertion sites in a difficult to lyse bacterium: *Lactobacillus casei*. (in press).

Licandro-Seraut H, **Scornec H**, Pédrón T, Cavin J-F, Sansonetti PJ. **2014**. Functional genomics of *Lactobacillus casei* establishment in the gut. Proc Natl Acad Sci USA 111:E3101-3109.

Perpetuini G, **Scornec H**, Tofalo R, Serrero P, Schirone M, Suzzi G, Corsetti A, Cavin J-F, Licandro-Seraut H. **2013**. Identification of critical genes for growth in olive brine by transposon mutagenesis of *Lactobacillus pentosus* C11. Appl Environ Microbiol. doi: 10.1128/AEM.01159-13.

Licandro-Seraut H, Brinster S, Van De Guchte M, **Scornec H**, Maguin E, Sansonetti PJ, Cavin J-F, Serrero P. **2012**. Development of an efficient *in vivo* system P<sub>junc</sub>-TpaseIS1223 for random transposon mutagenesis of *Lactobacillus casei*. Appl Environ Microbiol. doi:10.1128/AEM.00531-12.

### Communications orales

Licandro-Seraut H, **Scornec H**, Cavin J-F, Pédrón T, Sansonetti PJ. **2014**. Towards cellular microbiology of bacterial symbiosis: functional genomics of *Lactobacillus casei*. Closing Meeting HOMEOPITH ERC ADVANCED GRANT. Paris.

**Scornec H**, Licandro-Seraut H, Cavin J-F, Pédrón T, Sansonetti PJ. **2012**. Développement de la Signature Tagged Mutagenesis chez *Lactobacillus casei* pour identifier les gènes impliqués dans le maintien dans l'intestin. Workshop Interactions des microorganismes avec leurs environnements : circulation, adaptation. Dijon.

Licandro-Seraut H, **Scornec H**, Cavin J-F, Pédrón T, Sansonetti PJ. **2012**. Mutagenesis of the *Lactobacillus casei* genome: from library construction to phenotypic screening for colonization and beyond... HOMEOPITH ERC ADVANCED GRANT. Paris.

### Communications affichées

**Scornec H**, Licandro-Seraut H, Tichit M, Bouchier C, Pédrón T, Cavin J-F, Sansonetti PJ. **2014**. Identifying genome-wide transposon insertion sites to reassemble a random mutant library of *Lactobacillus casei* and maximize its use. EMBO conference. "Microbiology after the genomics revolution". Institut Pasteur, France.

Perpetuini G, **Scornec H**, Tofalo R, Serrero P, Schirone M, Suzzi G, Corsetti A, Cavin J-F, Licandro-Seraut H. **2013**. *Lactobacillus pentosus* genes involved in olive brine stress response. 5<sup>th</sup> FEMS Microbiology Congress. Leipzig, Allemagne.

**Scornec H**, Tichit M, Bouchier C, Cavin J-F, Pédrón T, Sansonetti PJ, Licandro-Seraut H. **2013**. Comment optimiser l'utilisation d'une banque de mutants par séquençage ? Forum des Jeunes Chercheurs. Dijon.

Licandro-Seraut H, **Scornec H**, Cavin J-F, Pédrón T, Sansonetti PJ. **2012**. Identification of crucial genes for *Lactobacillus casei* ileal persistence by signature tagged mutagenesis strategy. Tornado-FP7 Meeting. Annecy.

**Scornec H**, Licandro-Seraut H, Cavin J-F, Pédrón T, Sansonetti PJ. **2012**. Comment identifier les gènes de *Lactobacillus casei* qui lui permettent de se maintenir dans l'intestin ? Forum des Jeunes Chercheurs. Besançon.

Licandro-Seraut H, **Scornec H**, Cavin J-F, Pédrón T, Sansonetti PJ. **2012**. Développement de la Signature Tagged Mutagenesis chez *Lactobacillus casei* pour identifier les gènes impliqués dans le maintien dans l'intestin. Workshop Interactions des microorganismes avec leurs environnements : circulation, adaptation. Dijon.

Licandro-Seraut H, **Scornec H**, Cavin J-F, Pédrón T, Sansonetti PJ. **2012**. Development of Signature Tagged Mutagenesis of *Lactobacillus casei* to identify genes implied in gut persistence. The New Microbiology - Interacademies Symposium. Institut de France, Paris.

Licandro-Seraut H, **Scornec H**, Lee Hy, Lee Wj, Sansonetti PJ, Cavin J-F. **2011**. Drastic fluorescence instability in the probiotic *Lactobacillus plantarum* expressing GFP. 1<sup>st</sup> International CIGR Workshop on Food Safety. Dijon.

# **Avant-propos**

Au sein du laboratoire GPMA de l'ENSBANA jusqu'à fin 2011, puis au laboratoire PMB (Procédés Microbiologiques et Biotechnologiques) de l'UMR A 02102 PAM (Procédés Alimentaires et Microbiologiques) d'AgroSup Dijon/Université de Bourgogne, plusieurs thématiques sont axées sur la caractérisation de la réponse adaptative des microorganismes aux stress environnementaux. Chez les microorganismes en contact direct avec leur milieu, la transcription des gènes et la synthèse des protéines sont régulées de manière efficace à chaque changement des paramètres environnementaux (modification de la température, du pH, de la disponibilité en nutriments, de la concentration en sels, de la présence de composés toxiques) afin de permettre la survie cellulaire. La mise en place de cette réponse entraîne une diminution du taux de croissance ou une diminution de la survie des cellules. La régulation au niveau transcriptionnel est l'un des mécanismes fondamentaux de la réponse adaptative. Plusieurs systèmes de régulation existent pour aboutir à l'action du régulateur transcriptionnel en tant qu'activateur ou répresseur de l'expression d'un gène donné en se fixant au niveau de sa région promotrice comme nous le verrons dans la partie « Synthèse bibliographique » et ce sont les signaux environnementaux qui conditionnent son activité au niveau du promoteur.

Parmi les nombreux stress auxquels un microorganisme peut être soumis, le milieu intestinal semble être un milieu stressant. Afin de survivre dans cet environnement, le microorganisme va devoir résister à la présence d'un pH assez faible en sortie d'estomac, à la présence de sels biliaires, à la diminution en oxygène ainsi qu'aux défenses immunitaires de l'hôte. Si le microorganisme est un probiotique, il devra aussi être capable de s'établir et de persister dans l'intestin. Les mécanismes d'interactions symbiotiques sont encore peu connus et sont probablement liés aux interactions (liaisons, sécrétions...) à la surface des bactéries et des cellules hôtes. *Lactobacillus casei*, bactérie à Gram positif, commensale et probiotique, possède environ 330 gènes prédits d'après l'analyse du génome comme intervenant dans le processus cellulaire. Ces gènes pourraient potentiellement être impliqués dans cette symbiose. Dans ces 330 gènes, 207 sont impliqués dans la biogenèse de l'enveloppe cellulaire, 13 dans la motilité cellulaire et la sécrétion, 19 dans le trafic intracellulaire et la sécrétion, et 91 dans les mécanismes de transduction des signaux. Ces derniers vont, par conséquent, pouvoir moduler l'expression de gènes via des régulateurs de transcription afin d'adapter la bactérie à son environnement.

Afin d'avoir une vue globale de la totalité des gènes qui interviennent dans l'établissement de *L. casei* dans l'intestin, une approche de génétique inverse globale incluant un système de criblage efficace doit être mis en œuvre. Les travaux du stage de Master Recherche et de cette thèse s'inscrivent dans le projet ERC « HOMEOPITH » : "Homeostasis and rupture of the gut epithelium in the presence of commensals and pathogens". Ces travaux ont commencé sur la fin de la validation du système de mutagenèse par transposition chez *L. casei*. Ainsi les objectifs principaux du travail de thèse ont été d'identifier les gènes impliqués lors de

l'établissement *L. casei* dans l'intestin puis de caractériser fonctionnellement un opéron identifié. La thèse a bénéficié d'une allocation Ministère de l'Enseignement Supérieur et de la Recherche.

Ce travail est présenté en plusieurs parties. La partie « **Synthèse bibliographique** » porte sur les connaissances acquises dans cette thématique pendant la réalisation de cette thèse. Cette synthèse porte sur l'écologie intestinale, les mécanismes d'adaptation par la régulation transcriptionnelle et plus particulièrement sur les systèmes à deux composants puis sur l'architecture et le rôle de la paroi cellulaire dans les interactions hôte-microbe. La partie « **Contexte, objectif et stratégies** » permet de situer mon sujet de thèse dans la thématique et de présenter brièvement les objectifs et stratégies mises en œuvre au cours de mon travail de thèse. La partie « **Matériels et Méthodes** » rassemble les stratégies et techniques mises en œuvre dans ce travail. La partie « **Résultats** » reporte les résultats auxquels j'ai apporté ma contribution ainsi que ceux obtenus au cours de mon travail de thèse.

Cette partie est présentée sous forme de publications ou projets de publications précédés d'un résumé en français des principaux résultats. La partie « **Conclusion générale** » est une discussion des principaux résultats obtenus dans ce travail et par notre groupe de recherche. La partie « **Références bibliographiques** » correspond aux références des publications utilisées pour ce travail. Enfin, la partie « **Annexes** » présente deux articles, un publié et un en préparation, qui n'ont pas de lien direct avec mon travail de thèse mais dans lesquels j'ai pu m'impliquer.

# **Synthèse bibliographique**

## I. Commensales -Écologie intestinale

### I.1. Le tractus gastro-intestinal

Le tractus gastro-intestinal (TGI) est un écosystème complexe basé sur un équilibre entre l'épithélium gastro-intestinal, le système immunitaire et la flore intestinale (McCracken and Lorenz 2001). Ces trois composants sont essentiels pour un fonctionnement complet et un développement mature de l'écosystème gastro-intestinal. Le TGI a une superficie de plus de 300 m<sup>2</sup>, il représente la plus grande surface en contact avec l'environnement. Il permet la digestion et l'assimilation des aliments mais protège aussi des organismes pathogènes ; en effet, il constitue le premier organe immunitaire de l'organisme. Le TGI est composé de trois régions principales qui sont l'estomac, l'intestin grêle et le côlon (Figure 1).

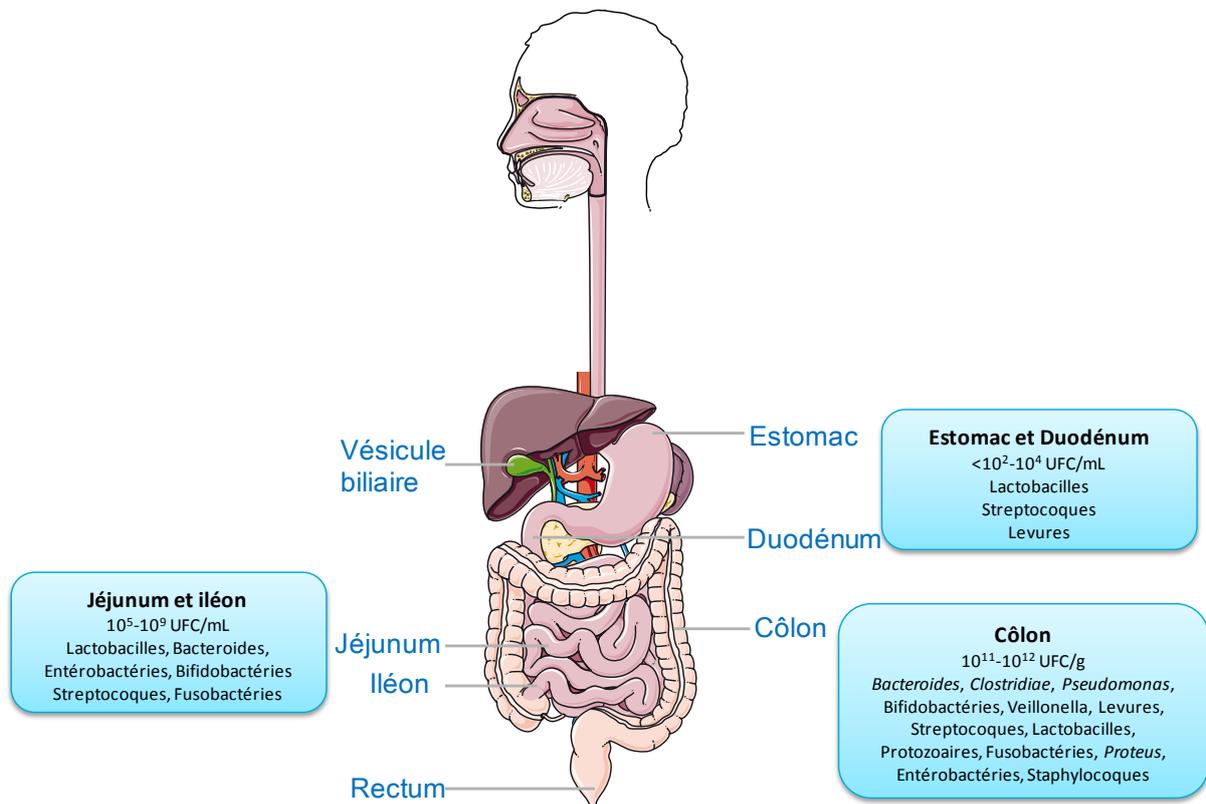


Figure 1 : Colonisation microbienne du tractus gastro-intestinal humain

### I.2. Composition de la flore du tractus gastro-intestinal

La composition en microorganisme du TGI est complexe et semble rester relativement stable chez l'adulte sain. La flore du TGI est composée de plus de 1000 espèces microbiennes distinctes. Cette microflore représente 10 à 20 fois le nombre total de cellules du corps

humain, soit  $10^{13}$ - $10^{14}$  cellules microbiennes (Hooper and Gordon 2001). La flore intestinale est constituée de bactéries résidentes dites bactéries commensales, endogènes et de bactéries exogènes issues de l'alimentation, de compléments alimentaires, ou probiotiques qui, elles, ne s'implanteront que de manière transitoire. Le microbiote intestinal est composé dans une très large majorité de bactéries anaérobies, la quantité d'*Archaea* et de *Fungi* est plus faible. Le microbiote est représenté en majorité par 5 phyla bactériens (Tremaroli and Backhed 2012) :

- Les *Firmicutes* incluant *Ruminococcus*, *Clostridium*, *Lactobacillus* (certaines souches sont probiotiques), et les producteurs de butyrate *Eubacterium*, *Faecalibacterium* et *Roseburia* ;
- Les *Bacteroidetes* incluant *Bacteroides*, *Prevotella* et *Xylanibacter* dégradant une variété de sucres complexes ;
- Les *Actinobacteria* incluant *Collinsella* et *Bifidobacterium* (qui contient des souches probiotiques) ;
- Les *Proteobacteria* incluant *Escherichia* et *Desulfovibrio* (qui contient des bactéries sulfito-réductrices) ;
- Les *Verrucomicrobia* incluant *Akkermansia* (qui sont spécialisés dans la dégradation du mucus).

La composition du microbiote est différente selon les régions du TGI (Fig. 1), ceci est en partie dû aux changements de conditions environnementales.

L'estomac est caractérisé par la présence d'acide mais aussi d'oxygène apporté par la déglutition. Son pH est compris entre 1,5 et 5. De ce fait, les microorganismes présents dans ce compartiment sont acidotolérants et anaérobies facultatifs tels que les lactobacilles, les streptocoques, les levures...

Dans le second compartiment, l'intestin grêle (ou petit intestin), composé du duodénum, jéjunum et iléon, les bactéries présentes sont les bactéries anaérobies facultatives comme les lactobacilles, les streptocoques, les entérobactéries et les bactéries anaérobies strictes notamment les bifidobactéries, les bactéroïdes et les clostridies.

Enfin dans le dernier compartiment, le côlon ou gros intestin, la flore microbienne est plus abondante. Elle est composée de bactéries anaérobies strictes telles que *Bacteroides*, *Clostridium*, *Bifidobacterium*... Les bactéries anaérobies facultatives comme les lactobacilles, les streptocoques, et les entérobactéries sont moins nombreuses.

La concentration microbienne dans ces différents compartiments est  $<10^2$  UFC/g dans l'estomac,  $10^2$ - $10^4$  UFC/g dans le duodénum  $10^5$ - $10^7$  UFC/g dans le jéjunum,  $10^7$ - $10^9$  UFC/g dans l'iléon et  $10^{11}$ - $10^{12}$  UFC/g dans le colon (Kleerebezem and Vaughan 2009).

Une étude métagénomique récente du microbiote intestinal (séquençage et analyse de l'ADN des microorganismes présents dans un échantillon) a révélé qu'à l'instar des groupes sanguins, chaque individu posséderait l'un des trois entérotypes (ou signatures microbiennes)

suyvants (Arumugam *et al.* 2011) : le premier dominé par les *Bacteroides*, le second par les *Ruminococcus*, le troisième par *Prevotella*. Ces entérotypes ne dépendraient pas de l'origine géographique, de l'âge ou de l'état de santé de l'individu.

La notion d'entérotypes est récente et d'autres analyses ont distingué seulement 2 entérotypes, l'entérotipe « *Ruminococcus* » étant inclus dans l'entérotipe « *Bacteroides* » (Wu *et al.* 2011).

### I.3. Fonctions/Rôle de la flore du tractus gastro-intestinal

Les interactions entre l'hôte et les microorganismes peuvent être de trois types : symbiose, commensalisme ou pathogénicité. **Le commensalisme** est une association étroite entre différentes espèces qui vivent de façon telle que l'une d'entre elles en tire bénéfice mais sans nuire aux autres espèces, au contraire du parasitisme. **La symbiose** est l'association étroite (cohabitation) entre deux organismes, chacun tirant un bénéfice de l'autre. Dans la relation symbiotique, les bactéries profitent d'une température stable, de l'oxygène et d'un apport nutritif, les hôtes tirent profit d'une capacité métabolique et/ou digestive plus importante et de l'exclusion compétitive des pathogènes (Collier-Hyams and Neish 2005). En considérant le génome global, la microflore représente environ 2 à 4.10<sup>6</sup> gènes. Les produits de ces gènes confèrent des capacités métaboliques que le génome de l'hôte ne possède pas (Hooper and Gordon 2001). Outre la fonction digestive, le microbiote intestinal a une fonction de protection, barrière, contre la colonisation du tube digestif par d'autres microorganismes qui seraient pathogènes. Ce microbiote permet aussi la stimulation du système immunitaire intestinal en le maintenant en état de veille.

Les lactobacilles sont présents tout au long du TGI et cela, dès la naissance par voie basse. En effet, les lactobacilles, présents dans la flore vaginale, sont transmis lors de l'accouchement. Ils peuvent rapidement coloniser l'intestin du nouveau-né qui est un environnement particulièrement permissif car stérile. *Lactobacillus casei* est un bacille Gram-positif, anaérobie facultatif, non mobile, non sporulé, hétérofermentaire et faisant partie des bactéries lactiques (LAB). *L. casei* est une bactérie alimentaire, elle peut être utilisée pour la fermentation des produits laitiers (yaourts, fromages...) et peut être retrouvée dans les olives. Comme d'autres LAB, *L. casei* est acido-tolérant et peut alors survivre au passage de l'estomac. Certaines souches de *L. casei* sont reconnues comme probiotiques (*L. casei* DN-114001 et Shirota) et sont alors beaucoup étudiées (Borruel *et al.* 2002; Borruel *et al.* 2003; Gionchetti *et al.* 2003). Les probiotiques sont définis par l'OMS comme étant des microorganismes vivants qui, lorsqu'ils sont ingérés en quantité suffisante, ont un effet bénéfique sur la santé de l'hôte. La souche *L. casei* ATCC 334 a été définie comme étant le néotype de l'espèce *L. casei* (Dicks *et al.* 1996). De plus, le génome de cette souche a été complètement séquencé et annoté (Cai *et al.* 2009). *L. casei* ATCC 334 a alors été choisi comme modèle pour cette étude.

## II. Mécanismes d'adaptation par la régulation transcriptionnelle

Tous les types de cellules connues utilisent des systèmes de transduction du signal pour répondre à un événement intra ou extracellulaire. Le rôle de ces systèmes est de pouvoir s'adapter à un changement environnemental ou simplement maintenir l'homéostasie cellulaire (Cashin *et al.* 2006). Un régulateur transcriptionnel va modifier le taux de transcrite ce qui va changer le niveau d'expression d'un gène.

### II.1. Les mécanismes de la transcription

Afin de comprendre les mécanismes de régulation transcriptionnelle, il faut connaître la transcription, résultant de l'action d'un complexe enzymatique : l'ARN polymérase (RNAP) c'est-à-dire de la synthèse de l'acide ribonucléique (ARN) à partir d'ADN. Chez les bactéries, il n'existe qu'une seule ARN polymérase. Le core de cette enzyme est constitué de différentes sous-unités :  $\alpha_2, \beta, \beta', \omega$  (Figure 2). Les sous-unités  $\alpha$  permettent de maintenir à l'écart le brin non transcrit de l'ADN et constituent une base permettant l'assemblage des autres sous-unités. Les sous-unités  $\beta$  et  $\beta'$  assurent la liaison à l'ADN et contiennent le site catalytique. La sous-unité  $\omega$  participe, quant à elle, à la liaison de l'enzyme sur la matrice d'ADN. La transcription est réalisée en plusieurs étapes commençant par l'initiation puis l'élongation et enfin la terminaison.

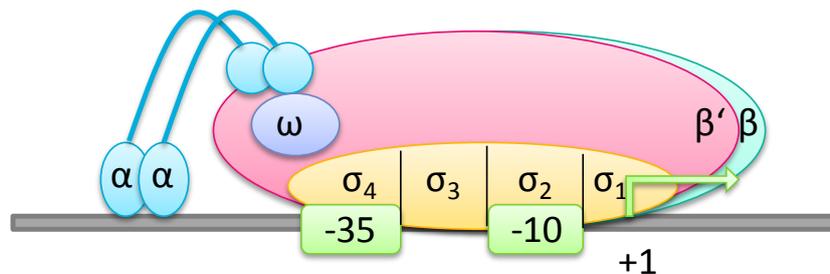
Au core de la RNAP va s'ajouter un facteur  $\sigma$  pour l'initiation de la transcription. Les facteurs  $\sigma$  sont des unités protéiques auxiliaires qui permettent d'augmenter l'affinité de la RNAP pour les régions promotrices : boîtes -35 (Site de connaissance) et -10 (Pribnow box). Chaque espèce bactérienne possède au moins un facteur  $\sigma$  de ménage, il fait typiquement partie de la famille des facteurs  $\sigma^{70}$ . Les facteurs  $\sigma$  ainsi que les facteurs de transcription plus spécifiques vont permettre de guider la RNAP pour contrôler l'expression de gènes cibles. A ce titre, ils constituent une première étape de régulation transcriptionnelle. Le complexe RNAP-facteur  $\sigma$  va former au départ un complexe fermé, ce qui signifie que l'ADN reste sous forme de double hélice. Les sous-unités 2 et 3 du facteur  $\sigma$  vont fixer les boîtes -10 et -35 respectivement. Puis la conformation de la RNAP est modifiée et forme alors un complexe ouvert, c'est-à-dire que la structure en double hélice se distord sur 10 à 15 paires de bases d'ADN dans la région d'initiation. Le primer d'initiation va alors être formé.

Une fois que 6 à 10 nucléotides sont assemblés, le facteur sigma se dissocie et permet ainsi d'abaisser l'affinité de la RNAP pour la région promotrice. La protéine NusA va alors se lier à la RNAP et permettre l'élongation.

La terminaison de la transcription peut se faire selon deux mécanismes : Rho-indépendante ou dépendante. Les terminateurs Rho-indépendants sont constitués d'une séquence inversée répétée suivie d'une série de TTTT (UUUU sur l'ARN transcrit). La tige boucle ainsi formée

va provoquer une pause de la RNAP. L'ARN transcrit ne sera alors apparié au brin d'ADN que par la séquence d'uridines. Les interactions A-U étant faibles, l'ARN transcrit se détache de la matrice. Les terminateurs Rho-dépendant sont, quant à eux, constitués d'une séquence consensus reconnue par la protéine Rho et va provoquer le détachement de l'ARN polymérase de l'ADN matrice.

L'analyse *in silico* du génome de *L. casei* montre la présence des facteurs  $\sigma$  RpoD, ComX,  $\sigma^{54}$  et  $\sigma^{24}$ .



**Figure 2 : Représentation schématique du complexe d'initiation de la transcription.**

La RNAP composée des sous-unités  $\alpha_2\beta\beta'\omega\sigma$  interagit avec les boîtes -35 et -10 de la région promotrice.

Chez les bactéries, un même promoteur peut réguler l'expression de plusieurs gènes regroupés. Ce regroupement de gène est appelé opéron. Dans le cas d'un opéron, les gènes sont transcrits en un ARN messenger unique. Chaque section de l'ARNm est ensuite traduite de façon différente. Ces ARNm sont dits polycistroniques car codent pour plus d'une chaîne polypeptidique. Les gènes d'un opéron encodent souvent plusieurs enzymes dans une même voie métabolique, ils sont souvent contrôlés de concert ce qui en fait un système très économique pour la cellule. Ils sont souvent soit tous transcrits soit aucun ne l'est.

## II.2. La régulation transcriptionnelle

Il existe trois grandes catégories de régulations transcriptionnelles :

- Les régulateurs transcriptionnels ou système à un composant ;
- Les systèmes à deux composants (TCS) ;
- Les systèmes à multiples composants.

Les systèmes à un composant permettent en majorité une réponse suite à la détection d'un signal cytoplasmique alors que les systèmes à deux ou plusieurs composants peuvent détecter des signaux extracellulaires. La particularité des systèmes à deux composants ou plus est que la régulation est plus fine.

### II.2.1. Les systèmes à un composant

Le système à un composant est le plus simple ; en effet, le signal va interagir directement avec le régulateur de transcription. Cette liaison du ligand va permettre un changement conformationnel du régulateur qui va pouvoir interagir directement à son tour avec la cible du promoteur et activer ou réprimer la transcription du gène. L'analyse sur 145 génomes procaryotes a montré qu'il y a plus de systèmes à un composant que de systèmes à deux composants environ 17 000 contre 4 000 respectivement (Ulrich *et al.* 2005). Dans le système à un composant, d'un point de vue structural, le régulateur transcriptionnel est composé d'un domaine de reconnaissance ou de liaison au signal dit « receveur » et d'un domaine de liaison à l'ADN dit « effecteur ». Différentes familles de régulateurs transcriptionnels existent ; ils sont en général classés par rapport au domaine de liaison à l'ADN. Le domaine majeur conservé pour 84% des régulateurs à un composant est constitué d'un motif HTH pour hélice-tour-hélice. Les autres motifs existants étant : HLH (hélice-boucle-hélice), doigt de zinc, feuillets  $\beta$  anti-parallèles...

Les régulateurs de transcription à un composant contenant le domaine HTH n'ont pas à 97% de domaine transmembranaire et sont alors prédits comme étant cytosoliques. La différence fondamentale entre le système à un composant et celui à deux est la détection du stimulus qui se ferait pour le premier intracellulaire et extracellulaire pour le deuxième (Ulrich *et al.* 2005).

La famille des régulateurs de transcription de type LysR (LTTR) est très bien caractérisée car très conservée et ubiquitaire chez les bactéries. Les LTTR sont impliqués dans la régulation de nombreux gènes qui codent pour des protéines intervenant dans différentes fonctions telles que le métabolisme, la division cellulaire, le quorum sensing, la virulence, la réponse au stress oxydatif...(Kovacikova and Skorupski 1999; Maddocks and Oyston 2008; Sperandio *et al.* 2007).

124 régulateurs de transcription ont été trouvés chez *L. casei* ATCC 334 ce qui correspond à environ 4,5% de son génome (Cai *et al.* 2009)

Dans les systèmes à deux composants ou plus, le domaine de liaison au ligand et le domaine de liaison à l'ADN se trouvent sur deux protéines séparées.

### II.2.2. Les systèmes à deux composants

Les systèmes à deux composants sont constitués d'une kinase membranaire qui agit comme senseur, c'est-à-dire qu'elle va répondre à un stimulus extracellulaire spécifique grâce au domaine de liaison au ligand. Le domaine de liaison à l'ADN sera, quant à lui, présent sur le second composant du système qui est le régulateur RR (Response Regulator) (West and Stock 2001) (Figure 3). Le senseur ou protéine histidine kinase (HK) s'auto-phosphoryle en réponse à un stimulus au niveau d'un résidu histidine spécifique. Ce phosphate est ensuite transféré

sur un résidu aspartate du régulateur cytosolique. La phosphorylation du régulateur entraîne une réponse adaptée dépendante du domaine effecteur. Ce domaine pouvant correspondre à un domaine de liaison à l'ADN (majeure partie), de liaison à l'ARN, enzymatique ou de liaison à des protéines (Galperin 2006). La constitution de chaque composant de ce système ainsi que leur mode d'action et leur implication sera d'avantage détaillé dans la partie II.3.

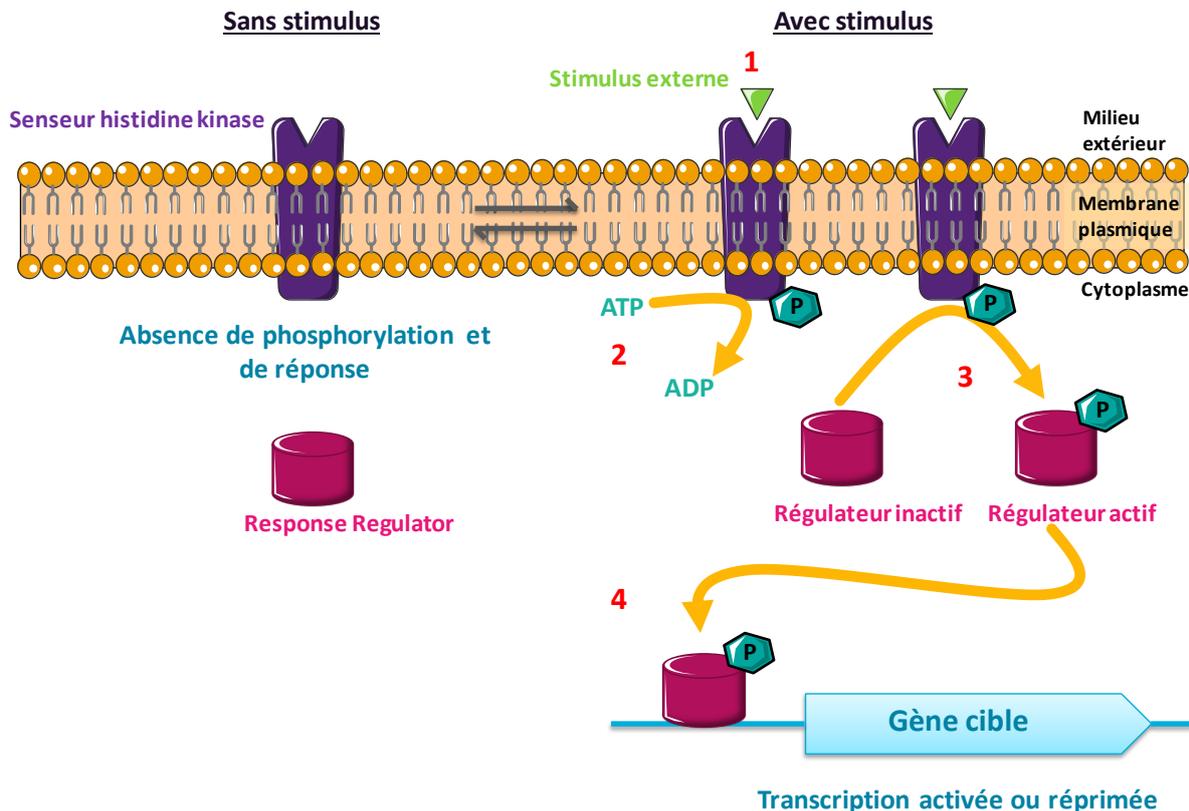


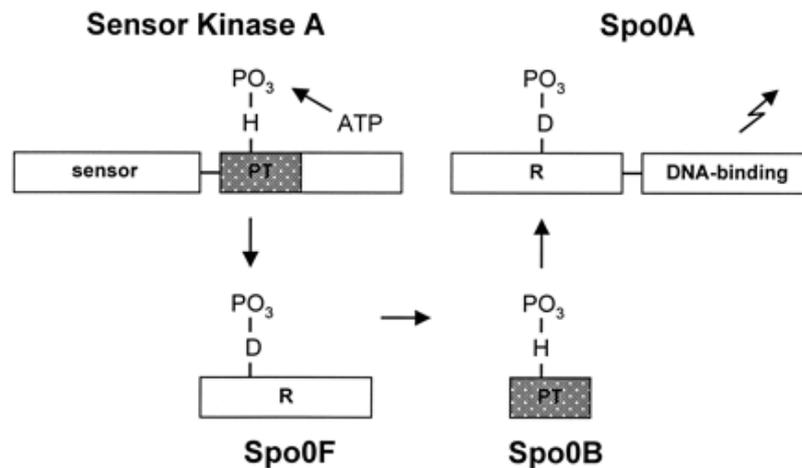
Figure 3 : Représentation schématique d'un système à deux composants.

(1) Détection d'un stimulus externe ; (2) Autophosphorylation du senseur en utilisant de l'ATP ; (3) Transfert du phosphate du senseur vers le régulateur ; (4) Fixation du régulateur sur le promoteur du gène cible.

### II.2.3. Les systèmes à multiples composants

A la place d'un seul transfert de groupement phosphate, certains systèmes sont caractérisés par un transfert multiple connu sous le nom de phosphorelais. Ce transfert de groupement phosphate a lieu entre des résidus histidine et aspartate (His→Asp→His→Asp) (Hoch 2000). Les systèmes de phosphorelais sont utilisés pour des voies complexes telles que la sporulation et le contrôle du cycle cellulaire chez les bactéries et comme voie de signalisation chez les eucaryotes. Un exemple de système de phosphorelais le plus étudié chez les bactéries est la régulation de la sporulation de *Bacillus subtilis*. Ce système fait intervenir un senseur histidine kinase A appelé KinA. Le phosphate de KinA va ensuite être transféré sur le résidu Asp d'une protéine intermédiaire nommée Spo0F (Stephenson and Hoch 2002). Cette

dernière va ensuite donner son phosphate à la protéine Spo0B correspondant à une histidine phosphotransférase (HPT) (Perraud *et al.* 1999). L'étape finale de ce phosphorelais étant le transfert du phosphate de la HPT Spo0B vers le résidu Asp du RR Spo0A (Figure 4). Le RR Spo0A va permettre la régulation de la transcription de 121 gènes (30 gènes individuels et 24 opérons) (Molle *et al.* 2003).



**Figure 4 : Système de phosphorelais impliqué dans la sporulation de *B. subtilis*.**

*D'après* (Hoch and Varughese 2001).

Le phosphorelais RcsCBD contrôlant une variété importante de processus cellulaires comme la mobilité, la division cellulaire, la formation de biofilm et la pathogénicité chez les *Enterobacteriaceae*, est composé : d'une kinase hybride RcsC qui possède un domaine transmetteur (domaine KD) et un domaine receveur (domaine RD) ; d'une histidine transférase nommée RcsD (domaine HPT) et enfin d'un RR nommé RcsB (Huang *et al.* 2006; Jayaratne *et al.* 1993; Takeda *et al.* 2001). Dans ce système, RcsC détecte un stimulus, phosphoryle son domaine KD qui ensuite va donner son phosphate au domaine RD. Histidine transférase RcsD va ensuite récupérer le phosphate pour finalement le transférer sur le RR RcsB (Figure 5).

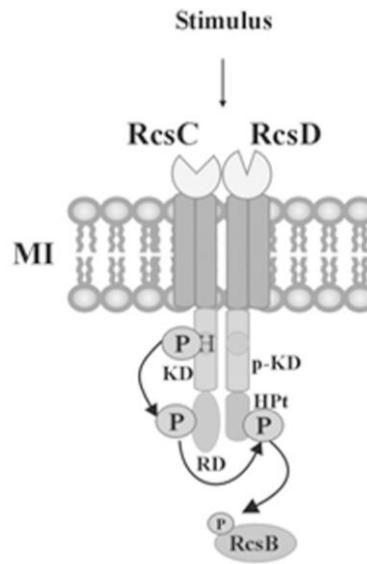


Figure 5 : Model de transduction du signal du système phosphorelay RcsCDB.

D'après (Pescaretti Mde *et al.* 2013).

### II.3. Les protéines d'un système à deux composants

Les systèmes à deux composants (TCS) permettent une réponse adaptée à une grande variété de stimuli. Les deux protéines possèdent chacune deux domaines spécifiques. L'HK possède un domaine N-terminal qui va sentir un stimulus spécifique par liaison ou réaction avec une molécule signal ou encore par interaction avec un stimulus physique. L'information est alors transmise et un changement conformationnel a lieu, il en résulte une activation du domaine transmetteur cytosolique C-terminal (Wolanin *et al.* 2002). Le transmetteur va activer le domaine N-terminal receveur du RR. Le RR donnera ensuite une réponse cellulaire adaptée grâce à son domaine effecteur C-terminal.

#### II.3.1. Le composant Histidine Kinase

##### Classification des HKs

La comparaison de séquences des TCS est très utilisée pour identifier un nombre de sous-familles. La classification des HKs est basée sur leur domaine transmetteur catalytique intracellulaire très conservé (Grebe and Stock 1999). Cette classification est basée sur six boîtes conservées dans le domaine transmetteur : **H**, **X**, **N**, **D**, **F** et **G**. Cette classification permet de mieux préciser les sous-groupes d'HKs. La composition générale des HKs est présentée dans la Figure 6.

La H-box contient l'histidine, site de phosphorylation. Une proline hautement conservée en position 5 après l'histidine acceptrice de phosphate est l'élément invariable des sous-familles

1 à 4. Dans les sous-familles 5 à 11, cette proline tend à disparaître. La distribution des résidus chargés et hydrophobes autour de l'histidine acceptrice de phosphate permet de différencier les différentes sous-familles. Les différentes boxes (boîtes) ainsi que leur spécificité sont présentées dans le Tableau 1.

Cependant cette classification néglige l'aspect fonctionnel du senseur et le processus de transduction du signal (Mascher *et al.* 2006). Regrouper les HKs par rapport à leur domaine senseur permettrait de les classer en fonction de leur rôle biologique qu'elles jouent dans la communication entre la cellule et son environnement. Néanmoins, la classification basée sur le domaine senseur N-terminal est problématique car ces domaines varient énormément à la fois au niveau de leur séquence, de leur topologie membranaire, de leur composition et de leur organisation. Par conséquent, une classification fonctionnelle ne peut être basée uniquement sur l'alignement des séquences mais requiert l'identification des domaines des hélices transmembranaires et la prédiction de l'agencement topologique de ces structures (Mascher *et al.* 2006). En se basant sur cette méthode, la majeure partie (si ce n'est pas toutes) des HKs peuvent être classés en trois groupes (Figure 7).

Le plus grand groupe correspond aux **HKs à perception extracellulaire de solutés ou de nutriments**. Ces HKs possèdent un domaine senseur extracellulaire qui est entouré de deux hélices transmembranaires, le domaine extracytoplasmique est composé de 50 à 300 résidus d'acides aminés. Le domaine kinase est quant à lui cytoplasmique.

Le second groupe contient les **HKs à perception membranaire**. Ces HKs peuvent sentir le signal membranaire grâce à la présence de 2 à 20 régions transmembranaires (TMR). Ces régions sont liées par de très courtes séquences intra ou extracellulaires (< 30 résidus d'acides aminés). Les HKs dans ce groupe peuvent être classées en six sous-groupes :

- Les kinases à petits senseurs (2 TMRs) impliquées dans les stress de l'enveloppe cellulaire et dans les processus de détoxification couplées à un ABC-transporteur ;
- Les thermosenseurs avec 4 ou 5 TMRs qui répondent à la fluidité membranaire ;
- Les senseurs globaux avec 6 TMRs ;
- Les HKs impliquées dans le quorum-sensing avec 6 à 10 TMRs ;
- Les HKs avec 12 à 20 TMRs qui montrent une homologie avec les protéines de transport et enfin ;
- Les HKs avec des domaines senseurs inconnus conservés avec 6 à 8 TMRs.

En somme, les stimuli détectés sont soit associés à la membrane, soit à l'interface membranaire. Les stimuli internes à la membrane incluent les propriétés mécaniques de l'enveloppe cellulaire (stress mécanique ou de turgescence) ainsi que les enzymes liées à la membrane, les gradients d'ions ou électrochimiques, les processus de transport ou la présence de composés affectant l'intégrité de la membrane.

Le dernier groupe regroupe les **HKs à perception cytoplasmique**. Ces HKs peuvent être cytoplasmiques ou avec une région transmembranaire, mais leur domaine senseur se trouve dans tous les cas dans le cytoplasme. Ces protéines vont détecter des solutés cytoplasmiques ou des protéines informant sur le métabolisme, l'état de développement ou le cycle cellulaire.

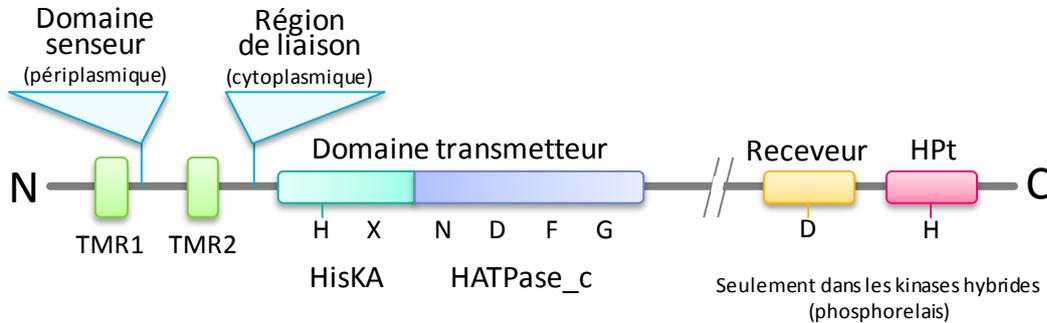
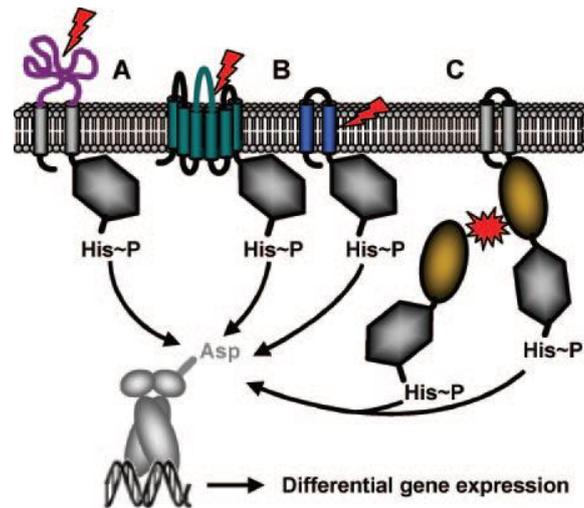


Figure 6 : Représentation des différents domaines constituant une histidine kinase.

Tableau 1 : Représentation des différentes boîtes retrouvées dans le domaine transmetteur des histidine kinases.

	H-box	X-box	N-box	D-box, F-box	G-box
HPK1a	+.Fh...SH-L+TPL..h		D...h..hh.NLh.NAh+ys	h.h.h.D.G.Gh hF..FR	G.GLGLshh..hh..HGG.h.h.G.Gs.F.h.h
HPK1b	+.FLA.MSH-h+TPL	S	D..+h.Qhh.NLh.NAhKFT	h.h.h.DsG.Gh hF..F.Q	G.GLGLshh..hh..MGG.h.h.G.Gs.F.h.h
HPK2a	F.DbAH-L+TPh..h		D..hh..hh.NLh.NAh+ys	h.h.h.D.G.G hF..FR	G.GLGLshh..hh..HGG.h.h.(G)(F)h.h
HPK2b	hh..hSH-hRTP.L.Rh		h..hh.NLh.NA.Ry	h.h.h.D.G.Ghs E hF.PFR	G.GLGLshh..hh..HGG.h.h.GGh.h.h.h
HPK3a	h..hTH.LKTP.h.L		Dh..hh.NLh.NAh+y	h.h.h.D.G.Gh hF.LRG.R	G.GLGLshh..hh..Y.G.h.h.Gh.h.h.h
HPK3b	h..F.ADhSHEL+VPh..h		h..hh.NLh.NA..y	h.h.h.DNG.Gh h..hF.RFy	G.GLGLshh..hh..HGG
HPK3c	h..h.H-hK.Ph..h		h..hh.NL..NAh.y	h.h.h.B.G.Gh hF..F	G.GLGLshh..hh.G.h.h.Gh.h.h.h
HPK3d	+.+hh..hSHDL.QPL..h		h..hh.NLh.NAh+yT	h.h.h.DTG.Gh hF..Fh	GhGLGLshh..hh..hGG.h.h.S..G.G..F.h.h
HPK3e	AAAHELG.TPL		+hh..h.NLh.NAVDyA	h.h.h.DDG.G..P.hh.RhG-PFh	GhGLGL.LLERSGA.h.F.N
HPK3f	h.H-L..h..h		h..hh.NLh.NAh.y	h.h.h.D.G.Gh h..hF..F	GhGLGL.hhE.HGG.h.h.Gs.F.h.h
HPK3g	AHHLNPh..h		h.QhNhNLh.NA.DA	h.h.h.D.G.Gh hFDPEFTTKP	GTGLGL.hhE.HGG.h.h.Gs.F.h.h
HPK3h	hSHDh.VPL..h		D...h..hh.NLh.NAhKF	h.h.h.D.G.Gh hF.hF.hH	GTGhGLshh..+hh..HGG.G.Gs.F.h.h
HPK3i	WhH.hKTP		KWL.Fhh.Qhh.NALKYS	T.h.h.D.G.Gh Dh.RhF..(F)	hGhLyLh..h..h..h.S.G.GT.h.h.F
HPK4	hG.hhs.hAHEh..Ph..h		D...h.QhNhNLh.NAhzh	h.h.h.D.G.Gh hF.PFTTK	G.GhGL.hh..hh..HGG.h.h.(S)Gs.F.h.h
Histidine kinase avec absence de proline dans la H-box					
HPK5	z.LR...HE..N		hh.hhGNLh..NAh.h	h.h.h.D.G.Gh hF..G	GhGL.hh..h.GG.h.h.Gs.F.h.h
HPK6	h.hh..hhRHPhhN		AB h..hh.NLh.NAh.HG	h.h.h.D.G.GhP hF..GF	G.GLGLyhh+hYyGG.h.h.G.F.h.h
HPK7	-..hA+hHD		h..hh.EAh.NAh+Hs	h.h.h.D.G.Gh Pas de F-box	GLGh..+h..hGG.h.h.G.h.h.h
HPK8	LQA...PHFLyN		h.hP.h.hO.hhENAh.y	h.h.h.D.G.Gh Pas de F-box	G.G.GLMhRh.FG.h.h.GT.h.h.h
HPK9	hFR.AH(S/T)KG.H.E		hh..h..PhhhhhRNA.DHG	hhh.h.DDG.Gh hF.PGFST	GRGhGhDVV+G.S.Gs.F.h.h
HPK 10	F+HDL.N		hh.h.R.h..hh.NAhE	Pas de D-box hF...FSTK	G.G.GLGL.F.h.h.h
HPK11	EhHHRh+NNLQ		TbhPh.hhh.E.Lhs.NAh+ys	h.h.h.DBG.Gh(F)	(G)shGLGG.G..F.h.F

Synthèse réalisée d'après (Grebe and Stock 1999). « . » représente des résidus variables ; « - » correspond aux résidus D ou E ; « + » pour K ou R ; « h » pour les résidus hydrophobes ; « y » pour les résidus structurellement similaires (H, F ou Y) ; « B » pour D ou N ; « z » pour E, D, N ou Q ; et « s » pour A, G, S ou T. Les parties surlignées en rose correspondent aux différences permettant de classer les différentes familles de HK. HPK : Histidine Protein Kinases.



**Figure 7 : Représentation schématique des trois groupes de HK.**

(A) HKs à perception extracellulaire, (B) HKs à perception membranaire, (C) HKs à perception cytoplasmique. Le stimulus est représenté par un éclair ou une étoile rouge. Les parties de la protéine impliquées dans la perception du signal sont colorées. D'après (Mascher *et al.* 2006).

### Composition des HKs

Les HKs peuvent être composés de nombreux domaines ; ne seront présentés ici que les domaines qui sont retrouvés dans les HKs de *L. casei*. La composition des HKs de *L. casei* est présentée dans la partie II.4.

#### Le domaine HAMP

Le domaine HAMP est localisé dans le cytoplasme, mesure environ 50 résidus d'acides aminés et correspond à deux hélices  $\alpha$  reliées par une boucle. Il ne participe pas à la détection du signal mais il s'agit d'un domaine important pour la transduction du système ; en effet, le signal perçu va entraîner la rotation des 4 hélices du domaine HAMP des protéines associées en dimère, ce qui va permettre l'activation du domaine transmetteur (Williams and Stewart 1999).

#### Le domaine PAS

Le domaine PAS (Per-Arnt-Sim) est localisé dans le cytoplasme, mesure environ 110 résidus d'acides aminés. Il est conservé et largement répandu dans les domaines senseurs à l'oxygène au potentiel redox et à la lumière. Il peut lier l'hème, la flavine et l'adénine (Taylor and Zhulin 1999).

Les domaines HAMP et PAS vont se trouver entre la dernière hélice transmembranaire et le domaine transmetteur.

#### Le domaine HisKA

Le domaine HisKA est localisé dans le cytoplasme, mesure environ 80 résidus d'acides aminés. Il correspond au domaine de dimérisation et au domaine phosphoaccepteur. Les H- et X-boxes vont se retrouver dans ce domaine.

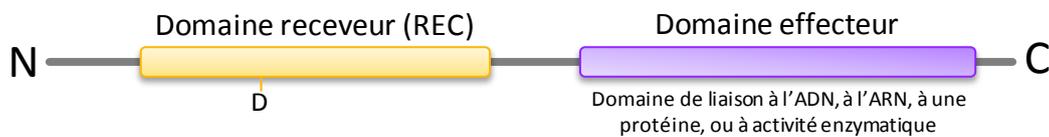
### Le domaine HATPase\_c

Le domaine HATPase\_c est aussi localisé dans le cytoplasme mesurant environ 140 résidus d'acides aminés. Il s'agit du domaine catalytique des HKs, il permet le phosphotransfert de l'ATP vers le domaine HisKA. Les N-, D-, F- et G-boxes vont se retrouver dans ce domaine.

Les domaines HisKA et HATPase\_c correspondent au domaine transmetteur de l'HK.

### *II.3.2. Le composant Response Regulator*

Les RRs sont composés en général de deux domaines : le domaine receveur REC et le domaine effecteur (Figure 8).



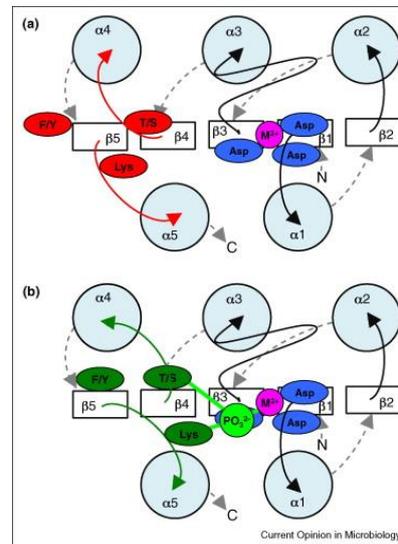
**Figure 8 : Représentation des différents domaines constituant un response regulator.**

#### Le domaine receveur

Le domaine receveur REC est un module commun à une grande variété de RR dans les systèmes de transduction du signal chez les bactéries. Ce domaine a été décrit pour la première fois en 1977 par Koshland lors de l'étude de la protéine CheY constituée uniquement du domaine REC (Koshland 1977). 17% des RRs bactériens ne possèdent que le domaine REC (Galperin 2010). Ces protéines contrôlent la motilité cellulaire (CheY), participent à la signalisation des phosphorelais (Spo0F) et, dans les interactions protéine-protéine, gouvernent le développement et la division cellulaire (DivK et CpdR) (Jenal and Galperin 2009).

Le domaine receveur N-terminal des RRs a trois grandes fonctions. La première est qu'il interagit avec l'HK phosphorylée et catalyse le transfert de phosphate vers son résidu aspartate. La seconde est de catalyser leur propre déphosphorylation et enfin la dernière est de réguler l'activité de leur domaine effecteur de manière phospho-dépendante (West and Stock 2001). Le domaine REC est composé d'environ 125 résidus qui sont organisés en 5 feuillets  $\beta$  entourés de 5 hélices  $\alpha$ . Chez la protéine CheY les résidus Asp12, Asp 13 et Asp57 participent à la coordination de l'ion  $Mg^{2+}$  requise pour le transfert du phosphate sur Asp57 du feuillet  $\beta_3$ , la Lys109 est aussi requise dans le site catalytique. Deux autres résidus sont aussi hautement conservés, le résidu Ser/Thr sur le feuillet  $\beta_4$  (Thr87) et le résidu Phe/Tyr sur le feuillet  $\beta_5$  (Tyr106). Ces deux résidus changent d'orientation en fonction de la phosphorylation de la protéine (Figure 9).

Des petites molécules comme l'acétyl-phosphate, le carbamoyl-phosphate, l'imidazole phosphate et le phosphoramidate peuvent servir de phospho-donneurs pour les RRs (Lukat *et al.* 1992)



**Figure 9 : Organisation spatiale schématique du domaine REC des RRs.**

Organisation avant (a) et après (b) phosphorylation. D'après (Bourret 2010).

Le problème de la classification des RRs en fonction de leur alignement de séquence est que deux RRs peuvent partager plus de 30% d'identité notamment dû à leur domaine REC alors qu'ils ne partagent pas du tout la même fonction (Galperin 2010). Afin de réaliser une classification fonctionnelle, les RRs sont regroupés par rapport à leur domaine effecteur qui correspond à un domaine de liaison à l'ADN, à l'ARN, aux protéines ou encore un domaine correspondant à une enzyme.

### Le domaine effecteur

#### *RR avec un domaine effecteur de liaison à l'ADN*

Une grande majorité (60%) des RRs contiennent un domaine effecteur de liaison à l'ADN. La plupart ont une structure tridimensionnelle correspondant à des variations du motif structurel de liaison à l'ADN : hélice-tour-hélice (HTH). La majeure partie de ces 60% de RRs avec un domaine de liaison à l'ADN peuvent être classés dans les familles **OmpR** (30%), **NarL** (17%), **NtrC** (9,1%) et **LytTR** (3%). Les deux motifs les plus communément retrouvés sont HTH et wHTH (motif ailé) (Figure 10).

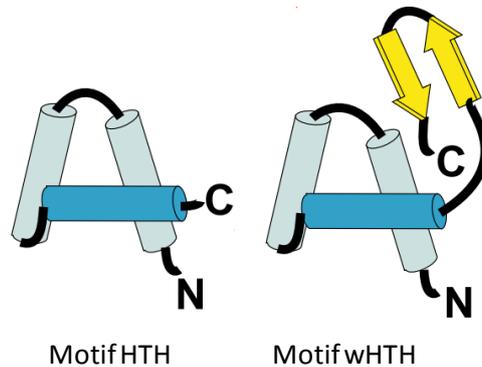
#### Le motif hélice-tour-hélice

La structure la plus simple du motif HTH est constituée de trois hélices. Les deux premières hélices se joignent par un tour de quatre résidus dont le deuxième est souvent une glycine. La deuxième hélice, qui s'insère dans le grand sillon de l'ADN, est nommée hélice de

reconnaissance. La troisième hélice, qui se trouve à l'avant dans une orientation horizontale, sert pour la stabilisation du motif.

### Le motif hélice ailée (winged helix, wHTH)

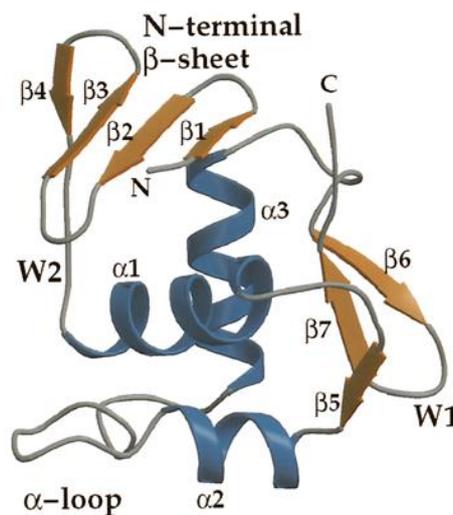
Le motif wHTH est une autre configuration du motif HTH de base. Ce domaine est caractérisé par la présence de deux feuillets  $\beta$  dans la région C-terminale. La reconnaissance du site est plus flexible grâce à cette aile ainsi formée.



**Figure 10 :** Représentation schématique des motifs HTH et wHTH de liaison à l'ADN.

Les feuillets sont représentés par des flèches jaunes, les hélices par des cylindres bleus, l'hélice de reconnaissance est représentée en bleu foncé. D'après (Aravind *et al.* 2005).

Les 98 résidus d'acides aminés composant le domaine effecteur des RR de la famille **OmpR** sont organisés en trois hélices  $\alpha$  entourées de deux modules de feuillets  $\beta$  antiparallèles constituant ainsi le motif wHTH. La topologie de ce domaine est  $\beta 1-\beta 2-\beta 3-\beta 4-\alpha 1-\beta 5-\alpha 2-\alpha 3-\beta 6-\beta 7$ . La région  $\alpha 2$ -tour- $\alpha 3$  forme le motif HTH. L'hélice  $\alpha 3$  correspond à l'hélice de reconnaissance qui va interagir avec l'ADN et l'hélice  $\alpha 2$  correspond à l'hélice de positionnement (Martinez-Hackert and Stock 1997) (Figure 11).



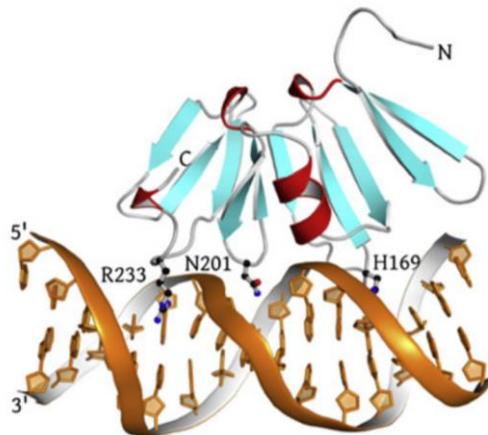
**Figure 11 :** Représentation tridimensionnelle du domaine effecteur d'OmpR d'*E. coli*.

D'après (Martinez-Hackert and Stock 1997).

Les 62 acides aminés formant le domaine effecteur des RR de la famille **NarL** d'*E. coli* sont quant à eux en un complexe de 4 hélices  $\alpha$ , les deux hélices centrales forment le motif HTH. Les deux autres hélices vont permettre l'assemblage en dimère. Lorsque NarL est non phosphorylé, le domaine effecteur est tourné vers le domaine receveur ce qui va empêcher la liaison à l'ADN. La phosphorylation de NarL sur son Asp59 N-terminal va entraîner une rotation de la partie C-terminale (effectrice) qui permettra alors la fixation à l'ADN (Baikalov *et al.* 1996).

Le domaine effecteur des RR de la famille **NtrC** est composé de deux domaines, un domaine ATPase de la classe AAA<sup>+</sup> de 230 résidus d'acides aminés, ce domaine permet de cliver l'ATP et de produire de l'énergie afin de former un complexe oligomérisé. Le second domaine de 80 résidus, Fis, pour Facteur de stimulation inversé, contient le motif HTH et va intervenir dans la dimérisation et l'oligomérisation. NtrC va interagir avec la  $\sigma^{54}$ -ARN polymérase (Yang *et al.* 2004).

Le domaine effecteur **LytTR** de la protéine AgrA de *Staphylococcus aureus* de 105 résidus ne comporte pas de motif HTH mais est un ensemble de 10 feuillets  $\beta$  organisé autour d'une hélice  $\alpha$ . Les RRs contenant ce motif sont généralement impliqués dans la régulation de facteurs de virulence ou dans la production de toxines (Sidote *et al.* 2008) (Figure 12).

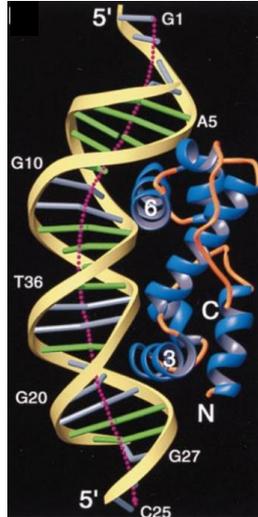


**Figure 12 : Représentation tridimensionnelle du complexe entre l'ADN et le domaine effecteur LytTR d'AgrA de *S. aureus*.**

Le domaine LytTR d'AgrA est composé de feuillets  $\beta$  en bleu et d'une hélice  $\alpha$  en rouge se positionnant sur l'ADN en orange. D'après (Sidote *et al.* 2008).

Le domaine effecteur **AraC-HTH** de 129 résidus de la protéine MarA d'*E. coli* est composé de sept hélices  $\alpha$  formant deux sous-domaines similaires contenant chacun le motif HTH. Il

possède deux hélices de reconnaissance, les hélices  $\alpha 3$  et  $\alpha 6$  (Figure 13). Ce domaine est présent en monomère en solution et ne contient pas de domaine de dimérisation (Rhee *et al.* 1998).

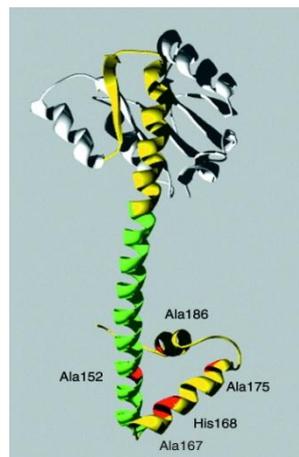


**Figure 13 : Représentation de la structure tridimensionnelle du complexe entre l'ADN et le domaine AraC-HTH de MarA d'*E. coli*.**

Le domaine AraC-HTH de MarA d'*E. coli* est composé de 7 hélices  $\alpha$  représentée en bleu. D'après (Rhee *et al.* 1998).

#### *RR avec un domaine effecteur de liaison à l'ARN*

Le plus connu des domaines de liaison à l'ARN est le domaine ANTAR pour AmiR and NasR Transcription Antitermination Regulators. Les régulateurs transcriptionnels de type AmiR et NasT possèdent ce domaine qui stimule la transcription en empêchant la terminaison de la transcription pour les terminateurs Rho-indépendants (Figure 14) (Galperin 2006).



**Figure 14 : Représentation de la structure tridimensionnelle de la protéine AmiR.**

Le domaine receveur est présenté en gris, la région représentée en jaune correspond au domaine ANTAR (résidus 101 à 190) avec la portion coiled coil importante pour la structure en dimère présentée en vert. Les résidus conservés correspondant à la signature du domaine ANTAR sont représentés en rouge (Shu and Zhulin 2002).

### *RR avec un domaine effecteur de liaison à une protéine*

La protéine CheV impliquée dans le chimiotaxisme est la protéine la plus étudiée qui possède un domaine effecteur de liaison à une protéine (CheW-like domain). Cette protéine est capable de fixer CheA comme le fait CheW ; elle peut, en plus de cela, être phosphorylée par CheA grâce à son domaine receveur REC (Alexander *et al.* 2010).

### *RR avec un domaine effecteur enzymatique*

Différentes activités enzymatiques peuvent être retrouvées dans des RRs. La protéine de chimiotactisme CheB possède un domaine REC en N-terminal et un domaine C-terminal méthylesterase. Les autres activités enzymatiques retrouvées chez d'autres RRs sont : l'activité diguanylate cyclase (RR de type PleD), c-di-GMP phosphodiesterase (RR de type VieA ou RpfG), phosphatase (RR de type RsbU) ou encore histidine kinase (Histidine kinase hybride) (RR de type RcsC). Le nombre de RRs possédant une de ces activités correspond à environ 1% du total des RRs analysés d'après (Galperin 2006).

### ***II.3.3. Transmission du signal de l'HK vers son RR et activité***

La transmission du signal des TCS est basée sur le transfert d'un phosphate du domaine transmetteur d'une HK vers le domaine receveur de son RR. L'HK est la protéine qui va détecter le stimulus grâce à son domaine senseur. Ce stimulus peut être extracellulaire, membranaire ou intracellulaire. La détection va ensuite faire intervenir le domaine transmetteur de l'HK. Ce domaine transmetteur possède deux sous-domaines : le domaine HisKA et le domaine HATPase\_c. Ce dernier correspond à la région catalytique de l'HK, c'est lui qui va permettre la phosphorylation à partir d'ATP. Le domaine HisKA va, quant à lui, permettre la dimérisation de l'HK. Ce domaine contient le site de phosphorylation, l'histidine de la H-box. Le phosphate de l'histidine de la H-box de la HK va ensuite être transféré sur le résidu aspartate conservé du domaine receveur REC du RR. Cette réaction va être catalysée par le domaine REC lui-même. Le RR possède deux domaines : le domaine REC et le domaine effecteur. Ce dernier domaine va ensuite être activé et va pouvoir exercer la régulation d'expression de gènes, la liaison à des protéines ou l'activité enzymatique... L'activation du RR peut se faire de deux manières différentes. Dans le premier cas, le domaine receveur inhibe le domaine effecteur et la phosphorylation permet de lever cette inhibition. Dans le deuxième cas, la phosphorylation du domaine REC va permettre la dimérisation ou l'oligomérisation du RR augmentant ainsi son affinité pour l'ADN ou pour des protéines. Certains RRs peuvent même combiner les deux cas (Stock *et al.* 2000). La réponse cellulaire du RR est proportionnelle à son degré d'activation.

## Mécanismes de régulation

Les HKs possèdent deux activités qui vont déterminer le niveau de phosphorylation des RRs. Premièrement, elles ont une activité d'autophosphorylation et secondement une activité phosphatase sur le RR. Les HKs ne possèdent pas toutes l'activité de phosphatase du RR. Dans ce cas, des protéines auxillaires peuvent effectuer la déphosphorylation du RR. Cependant de nombreux RRs ont une activité autophosphatase. Une autre façon de réguler le système est de réguler le taux de RRs lui-même. De nombreux systèmes à deux composants régulent leur propre système. Dans ce système, le RR est un activateur ou un répresseur de l'opéron codant pour les protéines du système à deux composants (Stock *et al.* 2000).

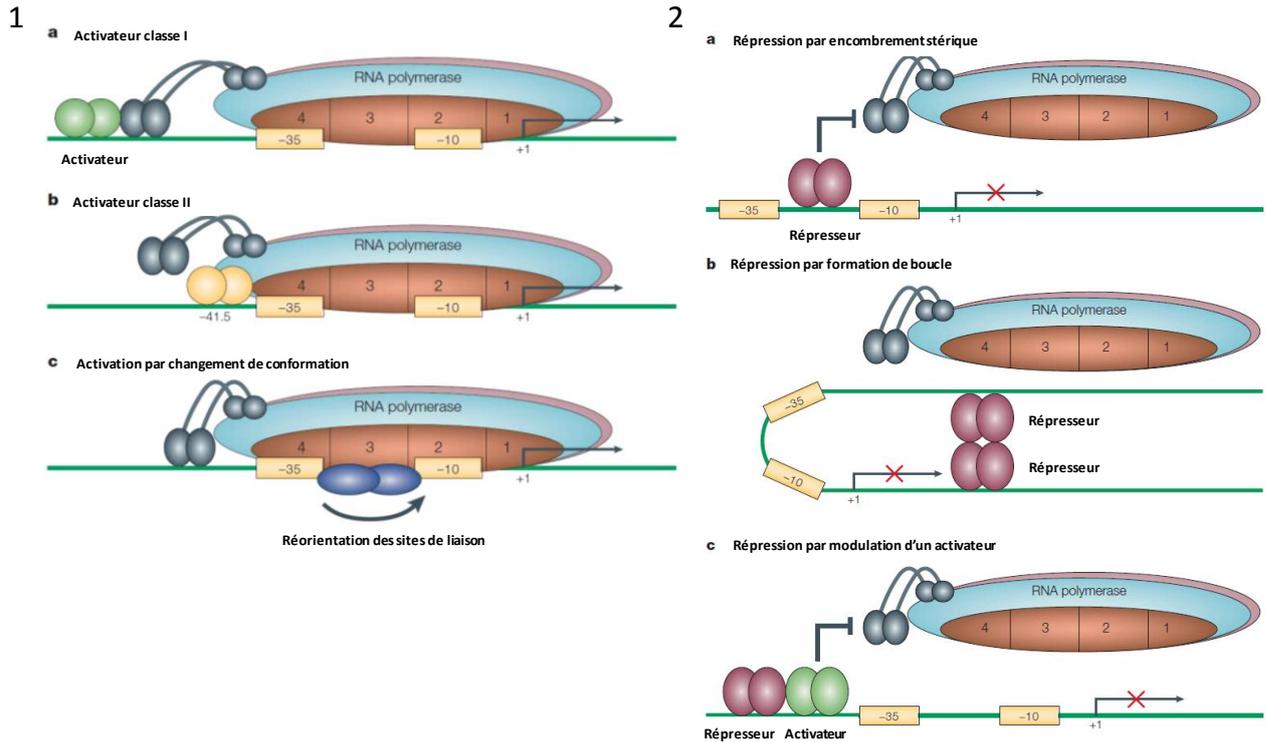
Les RRs qui ont un domaine effecteur de liaison à l'ADN sont soit des activateurs, soit des répresseurs de la transcription des gènes (Browning and Busby 2004) (Figure 15).

## Activateurs transcriptionnels

Dans le cas des activateurs de classe I, le RR peut se fixer bien en amont de la boîte -35, il interagit avec la partie C-terminale de sous unité  $\alpha$  de l'ARN polymérase, ce qui permet le recrutement de la polymérase sur le promoteur. Les activateurs de classe II se lient sur une région en amont mais proche de la boîte -35 (position -41.5) ; cette fois-ci l'activateur va interagir avec la sous-unité  $\sigma$  de l'ARN polymérase. Certains activateurs peuvent, quant à eux, venir se fixer entre les boîtes -35 et -10 ; l'activation de la transcription est due alors à un simple changement conformationnel de l'ADN, les boîtes -35 et -10 sont alors réalignées et l'ARN polymérase peut venir se fixer (Browning and Busby 2004).

## Répresseurs transcriptionnels

Les répresseurs inhibent la transcription en empêchant la fixation de l'ARN polymérase. Ils peuvent se fixer entre les boîtes -35 et -10 et réprimer la transcription grâce à leur encombrement stérique de la région promotrice. Ils peuvent aussi former une boucle au niveau des deux boîtes, ce qui empêche leur alignement ; dans ce cas, les répresseurs se positionnent en amont et en aval du promoteur et interagissent entre eux pour former la boucle. Enfin la dernière grande catégorie de répresseurs correspond aux régulateurs qui vont moduler l'activité des activateurs. Le répresseur va se lier à l'activateur qui ne pourra alors plus interagir avec l'ARN polymérase (Browning and Busby 2004).



**Figure 15 : Schématisation des différents mécanismes d'activation et de répression de l'initiation de la transcription par les régulateurs transcriptionnels.**

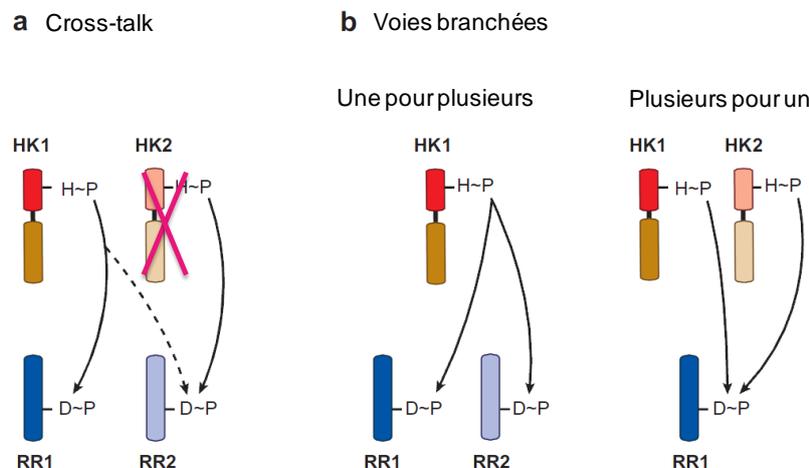
(1) mécanisme d'activation : 1a, activateur de classe I, 1b, activateur de classe II, 1c, activation par changement de conformation et réalignement des boîtes -10 et -35. (2) mécanisme de répression : 2a, répression par encombrement stérique, 2b, répression par formation de boucle, 2c, répression par modulation d'un activateur. D'après (Browning and Busby 2004).

### Cross-talk entre RRs et HKs

Les HKs et les RRs font partie chacun de familles de gènes paralogues et les membres de chaque famille partagent une importante homologie, à la fois au niveau de la séquence et de la structure. Cette similitude laisse à penser que les protéines de deux TCS de voies de signalisation différentes peuvent interagir entre eux (cross-talk) (Laub and Goulian 2007) (Figure 16). Cette interaction peut se produire chez *E. coli* entre les TCS VanS-VanR et PhoR-PhoB lorsque VanS-VanR sont exprimés de façon hétérologue. Le système VanS-VanR est impliqué dans la régulation de l'expression de gènes permettant la résistance à la vancomycine chez les entérocoques et d'autres bactéries Gram positif. Dans ce système VanS est une HK à perception membranaire possédant deux domaines transmembranaires et VanR est un RR de la famille OmpR (composé d'un domaine REC et d'un domaine trans\_reg\_C correspondant à un motif wHTH). Le système PhoR-PhoB est, quant à lui, impliqué dans la détection et la réponse à un changement de disponibilité en phosphate. Dans ce système, PhoR est une HK à perception extracellulaire et PhoB est un RR composé d'un domaine REC et d'un domaine trans\_reg\_C (wHTH). Une intercommunication entre la kinase VanS et le régulateur PhoB est possible chez *E.coli* mais seulement en absence de la kinase PhoR (Fisher

*et al.* 1995). De même, la kinase PhoR peut phosphoryler le RR VanR en absence de VanS (Haldimann *et al.* 1997).

D'autres systèmes de communications existent aussi entre des RRs et HKs de la même voie (voie branchée) (Figure 16b), une kinase peut parfois phosphoryler deux RRs. C'est le cas, par exemple, de CheA (HK impliquée dans le chimiotaxisme) qui est capable de phosphoryler deux RRs (une pour plusieurs), CheY et CheB (Kirby 2009). Inversement, un RR peut être phosphorylé par deux HKs (plusieurs pour un). Par exemple, trois HK hybrides nommés LuxN, LuxQ et CqsS peuvent phosphoryler une même phosphotransférase LuxU qui phosphoryle par la suite le RR LuxO (Henke and Bassler 2004).



**Figure 16 : Communication entre les différentes protéines des TCS.**

(a) cross-talk : communication entre des TCS de deux voies de signalisation différentes. (b) voies branchées : communications entre les protéines de la même voie de signalisation. HK : Histidine Kinase, RR : Response Regulator. (Laub and Goulian 2007).

#### II.4. Les systèmes à deux composants de *L. casei* ATCC 334

*L. casei* ATCC 334 a un génome de 2.95 Mb incluant son plasmide. Dans ce génome, 33 gènes codent 16 TCS complets et un incomplet (RR uniquement), faisant de *L. casei* un des lactobacilles possédant le plus de TCS (Zuniga *et al.* 2011). La répartition sur le génome de ces TCS est représentée dans la Figure 17. Les RRs font partie en majorité de la famille OmpR (12 RRs soit 70,6% des RRs), deux RRs (11,8%) appartiennent à la famille LytTR, deux autres RRs (11,8%) appartiennent à la famille NarL et enfin un RR est non classé. Certains de ces TCSs sont annotés d'après KEGG. Des mutants par insertion ont été obtenus chez *L. casei* BL23 pour chaque RR, ce qui indique qu'aucun des TCS n'est essentiel pour la croissance (Alcantara *et al.* 2011). Alcantara *et al.* ont testés ces mutants quant à leur capacité de croissance en MRS, en présence de bile 0,5% (m/v), de NaCl 0,6 M, à pH 3,75 et à 42°C, et quant à leur résistance aux antimicrobiens tels que la bacitracine, la vancomycine et la

nisine. Les résultats principaux de ce travail chez *L. casei* BL23 sont répertoriés dans le Tableau 2 par homologie de séquence des TCS de *L. casei* ATCC 334.

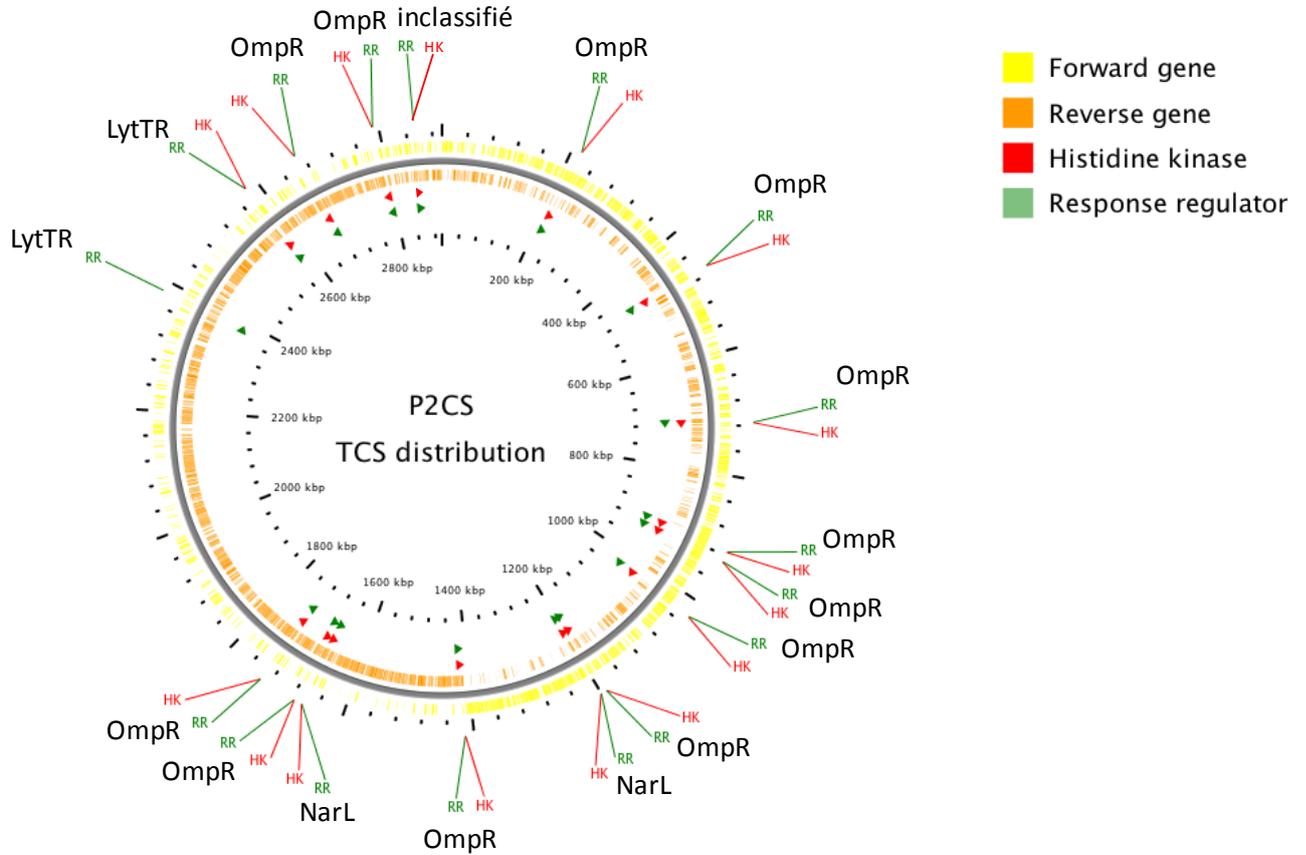
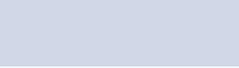


Figure 17 : Représentation de la position des TCSs sur le génome de *L. casei* ATCC 334.

Tableau 2 : Récapitulatif des caractéristiques des TCS de *L. casei* ATCC 334.

Nom des gènes codant pour le TCS	Représentation SMART des protéines composants le TCS	Nom courant du TCS	Caractéristique du TCS	Référence
LSEI_0219-0220			Mutant pour ce TCS est plus sensible à : Bile/NaCl/pH3,75 et à la bacitracine et nisine <b>Tolérance aux stress de la paroi</b>	Alcantara <i>et al.</i> 2011
LSEI_0461-0460			Augmentation de la transcription de LSEI_0460 lors d'un stress acide	Broadbent <i>et al.</i> 2010
LSEI_0711-0712		CiaR-CiaH	<b>Résistance au stress et compétence</b>	KEGG
LSEI_0934-0935		PhoP-PhoR	<b>Carrence en phosphate</b> Mutant pour ce TCS est plus résistant à la bacitracine et la nisine	KEGG Alcantara <i>et al.</i> 2011
LSEI_0950-0951				
LSEI_1041-1042			Mutant pour ce TCS est plus sensible à : Bile/NaCl/pH3,75/42°C Sensibilité nisine	Alcantara <i>et al.</i> 2011
LSEI_1209-1208				
LSEI_1223-1222		DesK-DesR	<b>Faible température</b>	KEGG
LSEI_1420-1419			<b>Tolérance aux stress de la paroi</b> Mutant pour ce TCS est plus sensible à la bacitracine et la nisine	Alcantara <i>et al.</i> 2011
LSEI_1666-1665			Mutant pour ce TCS est plus sensible à la bacitracine et la nisine	Alcantara <i>et al.</i> 2011
LSEI_1679-1678			Mutant pour ce TCS est plus sensible à la nisine et une croissance en MRS limitée par rapport à la souche parentale Transcription de LSEI_1679 diminuée lors d'un stress acide	Alcantara <i>et al.</i> 2011 Broadbent <i>et al.</i> 2010
LSEI_1740-1741			Mutant pour ce TCS est plus sensible à : Bile/pH3.75/42°C et la nisine <b>Tolérance aux stress de la paroi</b>	Alcantara <i>et al.</i> 2011
LSEI_2388-2389		AgrC-AgrA	<b>Quorum-sensing/bactériocine</b>	Kuo <i>et al.</i> 2013 Sturme <i>et al.</i> 2007
LSEI_2599-2600			TCS impliqué potentiellement dans le <b>Quorum sensing</b> d'après sa structure	Sturme <i>et al.</i> 2007
LSEI_2681-2680			Mutant pour ce TCS est plus résistant à la bacitracine	Alcantara <i>et al.</i> 2011
LSEI_2808-2807		WalR-WalK VicR/VicK YycF/YycG	<b>Métabolisme de la paroi</b> Mutant pour ce TCS est plus résistant à la bacitracine	KEGG Alcantara <i>et al.</i> 2011
LSEI_2868-2869		MaeK-MaeR CitA/CitB	<b>Métabolisme Malate/Citrate</b> Mutant pour ce TCS est plus résistant à la bacitracine Augmentation de la transcription de LSEI_2868 lors d'un stress acide	Landete <i>et al.</i> 2010 Broadbent <i>et al.</i> 2010

Les rectangles bleus représentent des régions transmembranaires, les rectangles roses correspondent à des régions de faibles complexités, les rectangles verts correspondent aux régions coiled-coil. Réalisé avec le logiciel SMART (<http://smart.embl-heidelberg.de/>) (Letunic *et al.* 2012; Schultz *et al.* 1998).

#### II.4.1. Détection des températures basses

Le TCS codé par les gènes LSEI\_1223-1222 correspond, d'après l'annotation du KEGG, au système DesK/DesR impliqué dans la détection et la réponse à de faibles températures. Chez *Bacillus subtilis*, DesR, le RR du système, régule l'expression du gène *des* encodant pour une  $\Delta 5$ -lipid desaturase. A une température de 37°C, DesK, l'HK du système, agit comme une phosphatase qui va déphosphoryler DesR, impliquant une diminution de la transcription. A température plus basse, la conformation des multiples domaines transmembranaires est altérée due à un changement de l'état lipidique de la membrane. Ce stimulus va permettre l'autophosphorylation de l'HK puis la phosphorylation du RR qui va pouvoir induire la transcription du gène *des* (Aguilar *et al.* 2001; Krell *et al.* 2010).

#### II.4.2. Métabolisme de la paroi

D'après l'analyse de la séquence protéique et les domaines fonctionnels, l'homologue le plus proche du TCS codé par les gènes LSEI\_2808-2807 est le système VicR/VicK, ou encore YycF/YycG. Ces systèmes sont à présent annotés WalR/WalK de par leur fonction ; en effet, ce TCS est impliqué dans le métabolisme de la paroi cellulaire lors de la phase exponentielle de croissance (Bisicchia *et al.* 2007). Ce système est impliqué dans la biosynthèse et le recyclage du peptidoglycane mais aussi dans la formation des biofilms chez *Staphylococcus aureus* (Dubrac *et al.* 2007). Il contrôle positivement la transcription des gènes codant pour les autolysines AltA et LytM. Une communication entre le système PhoRP et YycFG a été montrée chez *B. subtilis*, l'HK PhoR est capable de phosphoryler le RR YycF(WalR) lors d'une carence en phosphate (Howell *et al.* 2006). Le système WalR/WalK semble être essentiel pour les bactéries à Gram positif à faible GC% (Winkler and Hoch 2008). Cependant, l'inactivation de ce TCS chez *Lactobacillus sakei* face à de nombreux stress n'a pas montré de différence par rapport à la souche parentale (Morel-Deville *et al.* 1998).

#### II.4.3. Carence en phosphate

D'après l'analyse de la séquence protéique et les domaines fonctionnels, l'homologue le plus proche du TCS codé par les gènes LSEI\_0934-0935 est le système PhoP/PhoR. Ce système va permettre de détecter une carence en phosphate et va ainsi permettre, grâce à l'activation de la transcription de certains gènes, de récupérer de nouvelles sources de phosphate. Les cibles du RR (PhoP) de ce système sont les opérons codant pour les acides téichuroniques et les acides téichoïques. Ces derniers sont des polymères anioniques, ce sont des composants essentiels de l'enveloppe cellulaire et sont, de par leur composition, riches en phosphate. En absence de phosphate, *B. subtilis* remplace les acides téichoïques par les acides téichuroniques ne

contenant pas de phosphate, afin de restaurer le stock (Lahooti and Harwood 1999; Pragai *et al.* 2004).

#### II.4.4. Résistance au stress et compétence

D'après l'analyse de la séquence protéique et les domaines fonctionnels, l'homologue le plus proche du TCS codé par les gènes LSEI\_0711-0712 est le système CiaH/CiaR. Ce système a été caractérisé en majeure partie chez la bactérie pathogène *Streptococcus pneumoniae*. Ce TCS est impliqué dans de nombreuses fonctions telles que la compétence (capacité à recevoir de l'ADN étranger) (Dagkessamanskaia *et al.* 2004), la tolérance au stress oxydatif, la thermotolérance, la virulence (Quach *et al.* 2009). Il semble aussi être impliqué dans la biochimie de la paroi cellulaire des bactéries en effet, des mutations dans le gène codant pour CiaH (HK) confèrent une augmentation de la résistance aux bêta-lactames (antibiotiques agissant sur la paroi cellulaire) (Mascher *et al.* 2003; Zahner *et al.* 2002).

#### II.4.5. Quorum sensing-bactériocine

Le TCS incomplet ne possédant qu'un seul RR codé par le gène LSEI\_2389 est annoté par le KEGG comme faisant partie du système AgrC/AgrA impliqué dans le quorum sensing ou la production de bactériocine. Un pseudo-gène (LSEI\_2388) existe en amont de ce RR et le BLASTn de ce gène indique que la protéine codée semble correspondre à une histidine kinase. La régulation de la production de bactériocines telles que la sakacine chez *L. sakei*, ou la plantaricine chez *L. plantarum*, est généralement sous le contrôle d'un système à trois composants comprenant une HK, un RR de la famille LytR/AlgR/AgrA (Nikolskaya and Galperin 2002) et un petit peptide type phéromone appelé PAI (peptide auto-inducteur). Ce PAI est un peptide, comportant une double glycine dans sa séquence, qui est exporté dans le milieu extracellulaire grâce à un ABC-transporter. Le PAI peut lier le domaine extracellulaire de l'HK et par conséquent, activer le RR qui va augmenter la transcription de son propre cluster de gènes incluant le gène codant pour le PAI. En amont du TCS incomplet, se trouve des gènes codant pour un PAI putatif (LSEI\_2386), un ABC-transporter (LSEI\_2384) et une protéine accessoire pour la sécrétion de bactériocine (LSEI\_2381). Le gène de la HK est probablement annoté comme pseudo-gène car interrompu par une transposase (LSEI\_2387). L'action antimicrobienne du PAI a été testée sur différentes souches et il s'avère que ce peptide a une activité anti-*Listeria* (Kuo *et al.* 2013).

Le TCS codé par les gènes LSEI\_2599-2600 pourrait potentiellement être, lui aussi, impliqué dans le quorum sensing (QS-TCS). Le RR (codé par LSEI\_2599) de ce TCS possède un domaine LytTR et l'HK (codé par LSEI\_2600) fait partie de la sous-famille des HPK<sub>10</sub>. Ces particularités sont généralement retrouvées chez les QS-TCS (Sturme *et al.* 2007).

#### II.4.6. Utilisation de l'acide L-Malique

Le TCS codé par les gènes LSEI\_2868-2869 a récemment été caractérisé (Landete *et al.* 2010). Ce TCS est impliqué dans la régulation des gènes *maeP* (LSEI\_2867) et *maeE* (LSEI\_2866) requis pour la croissance en présence d'acide L-malique comme source de carbone. Le gène *maeE* code pour une malate oxydoréductase et le gène *maeP* pour un symporteur malate/Na<sup>+</sup>. Lorsque le gène *maeR* ou de *maeK* est délété, *L. casei* est incapable de pousser en présence d'acide L-malique et d'exprimer les gènes *maeE* et *maeP*. Ce TCS est aussi annoté comme intervenant dans la fermentation du citrate (CitA/CitB). *L. casei* possède, d'après le KEGG, les différentes protéines autres que le TCS intervenant dans ce métabolisme CitC (LSEI\_1860), CitD (LSEI\_1859), CitE (LSEI\_1858), CitF (LSEI\_1857), CitX (LSEI\_1856), CitG (LSEI\_1853).

#### II.4.7. Résistance aux peptides antimicrobiens

Les TCS codés par les gènes LSEI\_1420-1419 et LSEI\_1740-1741 semblent être impliqués dans la résistance aux peptides antimicrobiens (PAMs) (Revilla-Guarinos *et al.* 2013). Ces deux opérons ont été caractérisés chez *L. casei* BL23. Ils sont situés à côté de deux opérons codants pour des ABC transporteurs chez *L. casei* BL23 mais aussi chez ATCC 334 (LCABL\_16400-16410 ou LSEI\_1417-1418 et LCABL\_19580-19590 ou LSEI\_1738-1739) codant respectivement pour une peptide ABC transporter permease et pour une peptide ABC transporter ATPase. L'analyse transcriptionnelle a montré que le TCS codé par LCABL\_16420-16430 (soit LSEI\_1420-1419) est nécessaire pour l'induction de l'opéron de l'ABC transporter LCABL\_16400-16410 (soit LSEI\_1417-1418) en présence de nisine. Dans le cas de l'opéron LCABL\_19600-19610 (soit LSEI\_1740-1741), l'analyse a montré que les protéines codées par cet opéron n'ont pas de rôle direct dans la détoxification des PAMs comme c'est le cas pour les protéines du premier opéron. Les protéines codées par LCABL\_19600-19610 auraient plutôt un rôle sur le maintien des propriétés physicochimiques de l'enveloppe cellulaire en agissant sur la régulation (i) de l'opéron *dlt* LCABL\_08540-08580 (soit LSEI\_0793-0797); opéron impliqué dans la D-alanylation des acides lipotéichoïques nécessaire pour la résistance aux PAMs, (ii) d'un opéron ABC transporter isolé codé par LCABL\_21670-21680 (soit LSEI\_1993-1994) et (iii) du gène *mprF* (impliqué dans la L-lysinylation des phospholipides) (Figure 18).

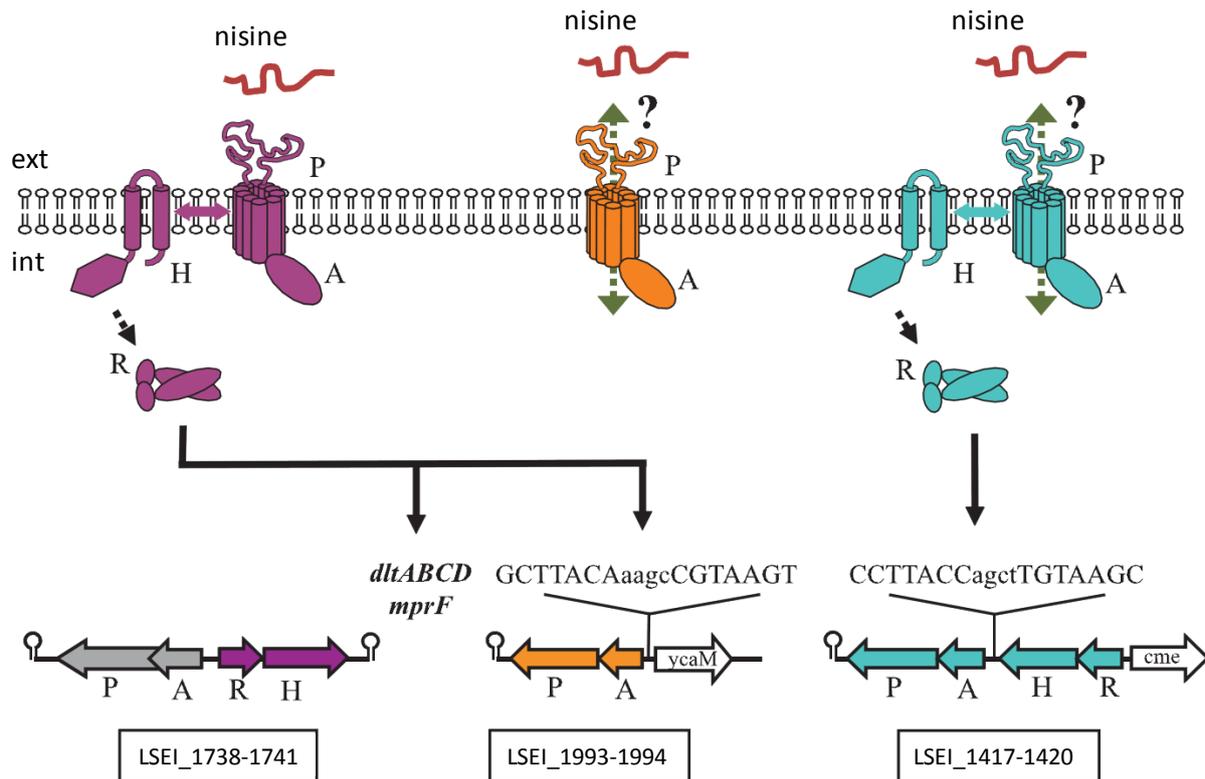


Figure 18 : Représentation des voies de signalisation senseur et de résistance de *L. casei* aux PAMs.

P : Permease, A : ATPase, H : Histidine kinase, R : Response Regulator. Les séquences putatives de fixation des régulateurs au niveau des promoteurs sont présentées en lettres majuscules. La nisine est représentée comme substrat interagissant avec les perméases. Les transports putatifs de la nisine sont présentés à travers les perméases par des flèches pointillées vertes. Le transfert de signal entre les transporteurs et les TCS est représenté par une flèche double entre la bicouche de phospholipides de la membrane. Les flèches noires pointillées représentent le phosphotransfert putatif entre HK et le RR. L'activation transcriptionnelle est représentée par des flèches noires.

#### II.4.8. Résistance au stress acide

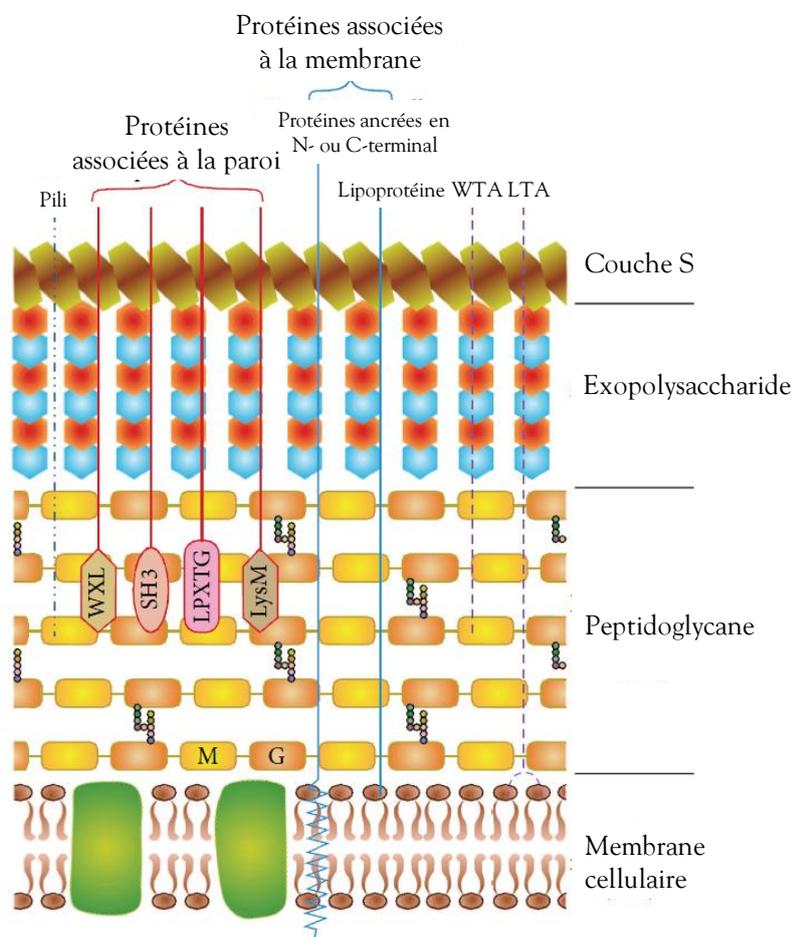
Les TCS codés par les gènes LSEI\_0461-0460, LSEI\_2868-2869 et LSEI\_1679-1678 semblent être impliqués dans la résistance au stress acide. En effet, une étude a récemment montré une augmentation de la transcription des deux premiers TCS ainsi qu'une diminution pour le dernier (Broadbent *et al.* 2010).

### III. Architecture et rôle de la surface cellulaire des Lactobacilles dans les interactions hôte-microbe

Les espèces du genre *Lactobacillus* peuvent exercer un effet favorable sur la santé dans le tractus gastro-intestinal grâce à différents mécanismes tels que l'inhibition des pathogènes, le maintien de l'équilibre microbien, l'immunomodulation et l'amélioration de la fonction de barrière. Les différentes espèces du genre *Lactobacillus* induisent des réponses différentes

chez l'hôte, ceci grâce à la diversité de l'architecture de la surface cellulaire et leur capacité à exprimer certains composants de surface ou excréter certains composés qui interagissent directement sur les cellules hôtes (Sengupta *et al.* 2013).

La paroi cellulaire des lactobacilles, bactéries à Gram positif, est composée d'un assemblage en plusieurs couches de polymère de peptidoglycane orné d'acides téichoïques (wall teichoic acids (WTA) ancrés au peptidoglycane et/ou lipoteichoic acids (LTA) ancré à la membrane), d'exopolysaccharides (EPS), de filaments protéiques appelés pili et de protéines associées à la paroi cellulaire. Certaines espèces de lactobacilles présentent une couche paracrystalline de protéines appelée couche S qui entoure le peptidoglycane (Figure 19).



**Figure 19 : Représentation schématique de l'enveloppe cellulaire des Lactobacilles.**

Les protéines associées à la paroi cellulaire sont attachées à la paroi de manière covalente (protéines LPxTG) ou non (protéines avec un domaine LysM, SH3, ou WXL). Les protéines associées à la membrane sont soit ancrées dans la membrane cellulaire (lipoprotéine), soit attachées grâce à une hélice transmembranaire N- ou C-terminale. G, N-acétylglucosamine ; M, acide N-acétylmuramique. D'après (Sengupta *et al.* 2013).

### III.1. Le peptidoglycane

#### III.1.1. Structure et composition du peptidoglycane

Le peptidoglycane (PG) est le composant majeur et essentiel de la paroi cellulaire des bactéries ; il détermine la morphologie et préserve l'intégrité de la cellule. La nature élastique du PG lui permet de résister à des forces d'étirement causées par exemple par la pression de turgescence, d'empêcher l'entrée de grosses molécules dans la cellule tout en permettant la sécrétion de grosses protéines. Le peptidoglycane est composé d'une partie glucidique (=polysaccharide) et d'une partie peptidique. Le polysaccharide du peptidoglycane est un polymère de glycosaminopeptide composé de *N*-acétyl-glucosamine (NAG) et d'acide *N*-acétyl-muramique (NAM) liés par des liaisons osidiques  $\beta$  1 $\rightarrow$ 4.

Les chaînes de polysaccharides sont liées entre elle par des ponts peptidiques entre deux NAMs formant la réticulation du peptidoglycane. En effet, chaque NAM possède un pentapeptide composé d'une alternance d'acides aminés L- et D-. Des variations considérables apparaissent dans la composition de base du pentapeptide, ce qui confère des caractéristiques spécifiques à la souche bactérienne (Asong *et al.* 2009). Ces variations sont présentées dans la (Figure 20). Ces pentapeptides sont ensuite liés entre eux et cet ensemble forme le pont peptidique. La liaison entre les pentapeptides est elle aussi variable (Vollmer *et al.* 2008). Les variations de cette liaison sont présentées dans la Figure 21. En somme, ces ponts peptidiques permettent l'élasticité du PG.

Les modifications du peptidoglycane telles que l'élimination du groupement acétyl du NAG (Araki *et al.* 1980), la O-acétylation du NAM ou du NAG (Bernard *et al.* 2011) ou encore la substitution du C6 du NAM par des acides téichoïques ou téichuroniques (Archibald 1985) peuvent survenir et affecter l'interaction avec l'hôte. En effet, ces modifications peuvent affecter la physiologie de la paroi cellulaire en augmentant la sensibilité à l'autolyse, la résistance au lysozyme et l'hydrophobicité de l'enveloppe cellulaire, affectant ainsi la reconnaissance par les récepteurs de l'hôte et l'adhésion des bactéries (Sengupta *et al.* 2013).

La structure du PG de *L. casei* BL23 a été déterminée, son peptide de base est composé de L-Ala, D-Gln, L-Lys, et D-Ala. Les résidus 3 et 4 sont liés par un pont peptidique composé d'un résidu D-Asn. Le NAM peut aussi être O-acétylé chez *L. casei* (Figure 22) (Regulski *et al.* 2012).

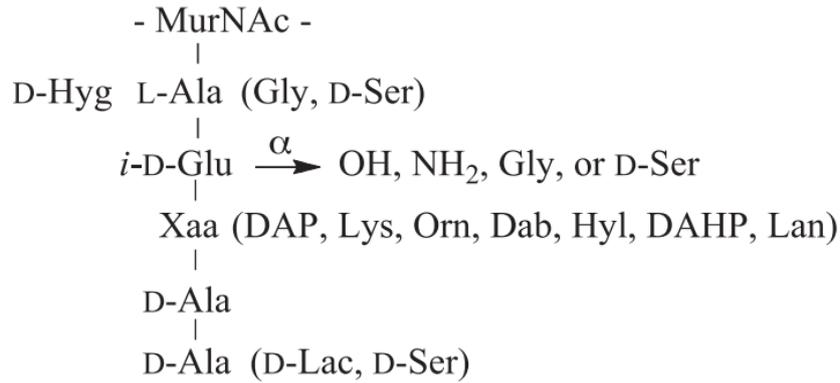


Figure 20 : Variation de la composition du pentapeptide du peptidoglycane.

Les résidus entre parenthèse peuvent remplacer l'acide aminé correspondant. L'acide  $\alpha$ -carboxylique de l'iso-D-Glu peut être modifié par un amide, un Gly ou D-Ser. (DAP, acide 2,4-diaminopimélique ; DAPH, acide 2,6-diamino-3-hydroxypimélique ; Hyg, acide thréo-3-hydroxyglutamique ; Hyl, hydroxylysine ; Lan, lanthionine ; Orn, ornithine). D'après (Asong *et al.* 2009).

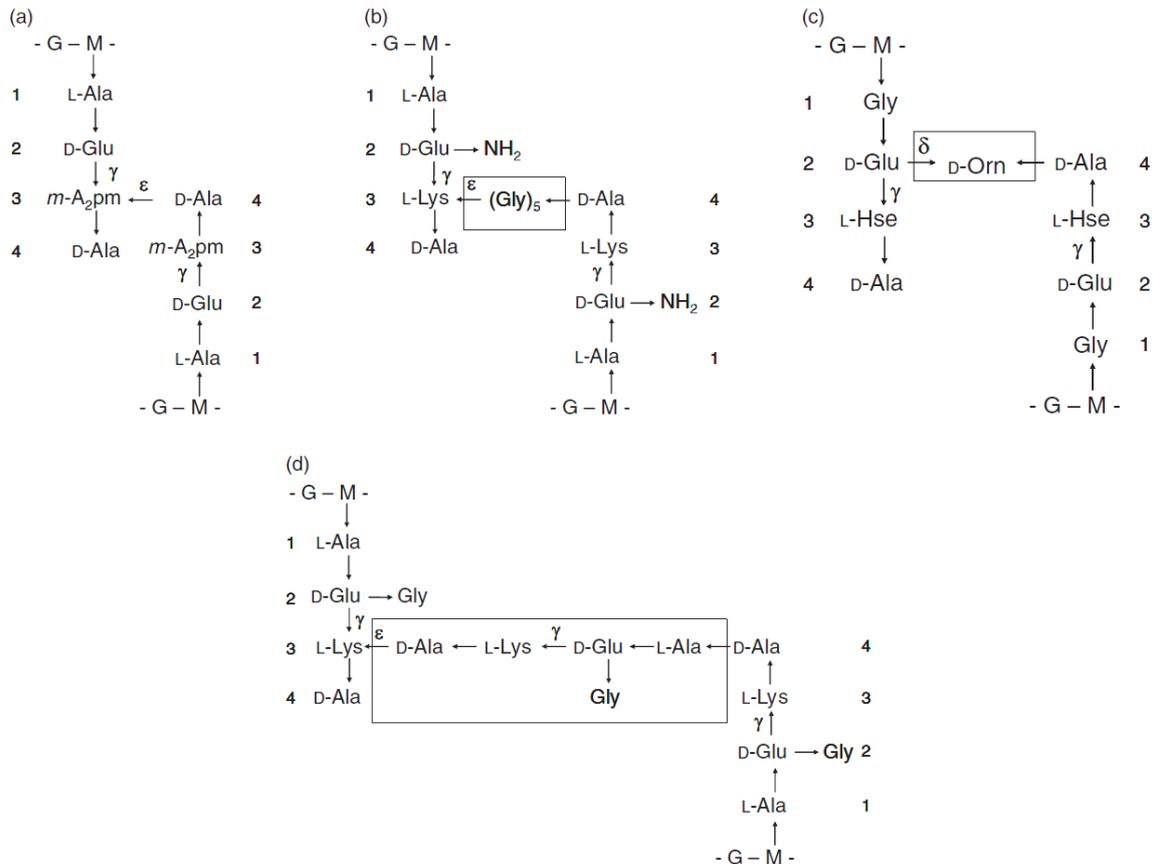


Figure 21 : Exemple de pont peptidique.

(a) *Escherichia coli* (pont direct entre les résidus 3 et 4) ; (b) *Staphylococcus aureus* (liaison des résidus 3 et 4 par un pont de pentaglycine) ; (c) *Corynebacterium pointsettiae* (liaison des résidus 2 et 4 par un pont avec une D-ornithine) ; (d) *Micrococcus luteus* (liaison des résidus 3 et 4 par un pont correspondant au peptide du NAM). G, N-acétylglucosamine ; M, acide N-acétylmuramique. D'après (Vollmer *et al.* 2008).

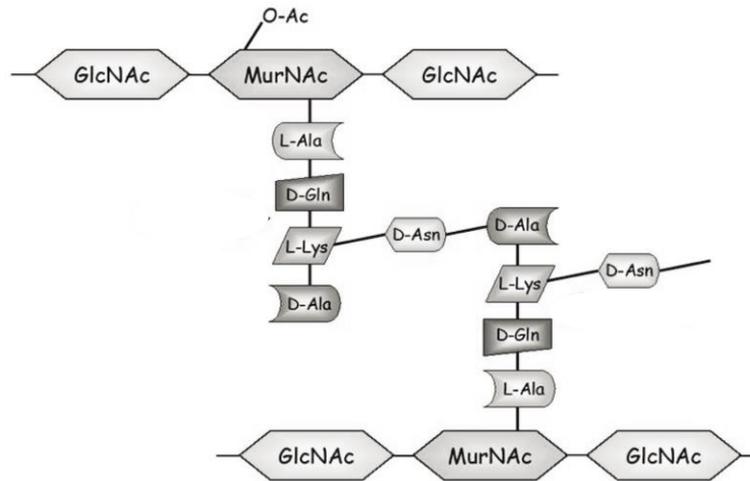


Figure 22 : Représentation de la structure du peptidoglycane de *L. casei* BL23.

GlcNAc, *N*-acétylglucosamine ; MurNAc, acide *N*-acétylmuramique. MurNAc peut être *O*-acétylé ou non. D'après (Regulski *et al.* 2012).

### III.1.2. Les Penicillin-Binding Proteins (PBPs)

Les PBPs sont des acyl-serine transférases. Elles sont impliquées dans l'assemblage, le maintien et la régulation des caractéristiques de la structure du peptidoglycane.

Les PBPs catalysent la polymérisation d'une chaîne glycanes (transglycosylation) et la liaison entre deux chaînes glycanes (transpeptidation). Certaines PBPs peuvent hydrolyser la dernière D-Alanine des pentapeptides (DD-carboxypeptidation) ou hydrolyser le pont peptidique entre deux chaînes glycanes (endopeptidation) (Figure 23). Par homologie avec leur substrat naturel, le pentapeptide (possédant une queue D-Ala-D-Ala), la pénicilline et les autres  $\beta$ -lactames sont des substrats de ces enzymes et peuvent agir comme inhibiteurs compétitifs (Strominger and Tipper 1965).

Les PBPs sont des protéines transmembranaires et leur site actif est accessible dans l'espace pariétal. Il existe différentes sortes de PBPs ; elles se différencient par leur faible poids moléculaire (LMW) ou haut poids moléculaire (HMW). Dans la plupart des cas, les PBPs de haut poids moléculaire ont une activité transglycosylase et/ou transpeptidase alors que les PBPs de faible poids moléculaire ont plutôt une activité carboxypeptidase et/ou endopeptidase (Ghuysen 1991).

Elles sont aussi organisées en classe A ou B pour les HMW, en fonction du domaine catalytique se trouvant dans la région N-terminale, la partie C-terminale correspondant à l'activité peptidase (Tableau 3). Les PBPs de classe A possèdent une activité glycosyltransférase, alors que celles de classe B possèdent un domaine N-terminal jouant un rôle dans la morphogénèse de la cellule et qui interagit avec d'autres protéines impliquées dans le cycle cellulaire.

Les LMW sont souvent caractérisées dans la classe C des PBP. Les PBP ont récemment été associées à d'autres fonctions telles que leur implication dans la virulence, la morphologie et leur interaction avec les cellules hôtes (Tableau 4). Ce tableau met en évidence un rôle plutôt tourné vers la virulence/pathogénicité pour les HMW et vers la morphologie de la cellule pour les LMW.

Trois groupes d'acides aminés (motifs) sont assez bien conservés dans les centres catalytiques des PBP, ces motifs sont SXXXK, S/YXN/C et [K, H] [T, S] G (Ghuysen 1991; Goffin and Ghuysen 1998).

Peu d'études portent sur le rôle de LMW PBP. A part leur implication dans le maintien de la morphologie et dans la résistance aux  $\beta$ -lactames, leur rôle reste encore peu connu.

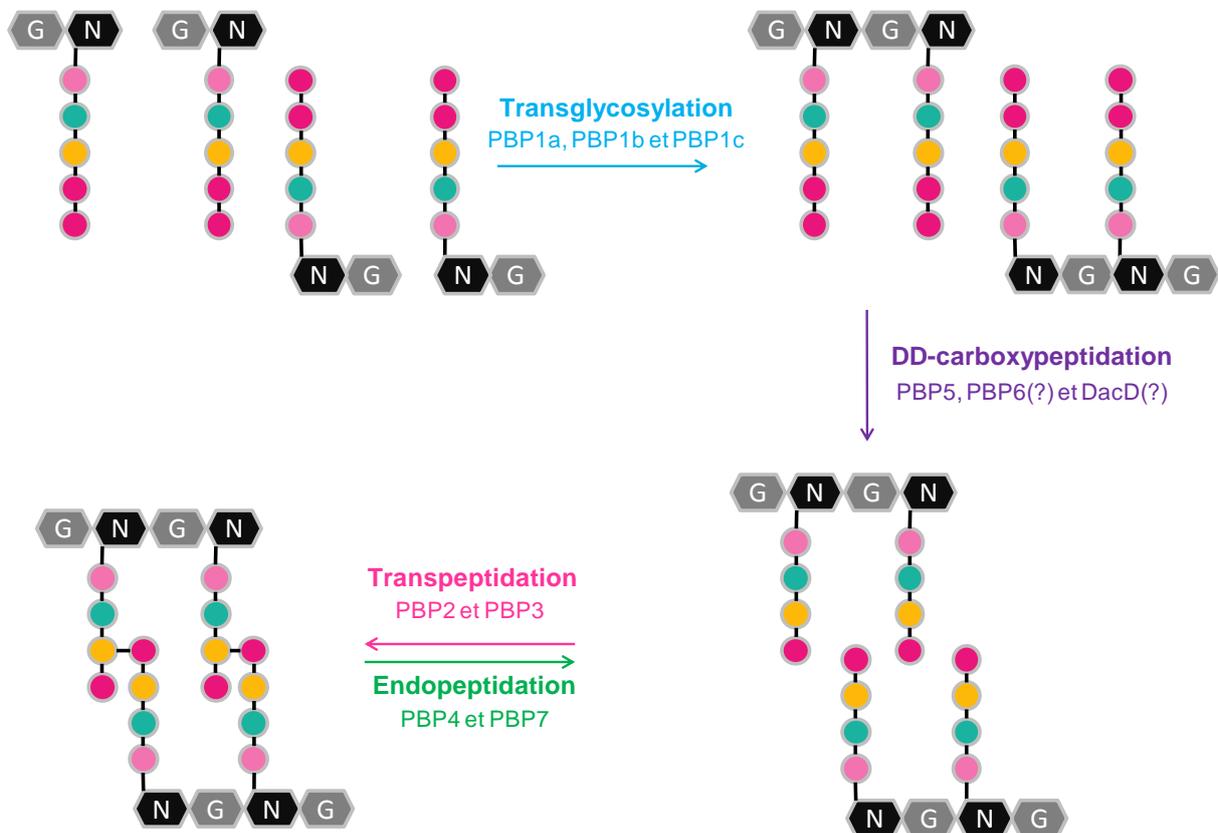


Figure 23 : Représentation des différentes fonctions des PBP dans la biosynthèse du peptidoglycane chez *E. coli*.

En bleu : les HMW PBP (PBP1a, 1b et 1c) polymérisent la chaîne de glycane en associant une molécule de NAG (N-acétylglucosamine) et une molécule de NAM (N-acétylmuramique). En violet : la PBP5 (éventuellement la PBP6 et DacD dont la fonction est incertaine) clive le dernier D-Ala de la chaîne peptidique du NAM. En rose : les PBP2 et PBP3 synthétisent la liaison entre deux brins de peptidoglycane (PG) (entre les résidus D-Ala et L-Lys ou l'acide diaminopimélique : DAP) par un pont direct. En vert : les PBP4 et PBP7 clivent la liaison entre deux brins de PG. Les points rose clair représentent le résidu L-Ala, en turquoise le résidu D-Glu, en orange le résidu DAP ou L-Lys, en rose foncé les résidus D-Ala.

Tableau 3 : Classification des PBP de sept bactéries.

	Class A							Class B							Class C											
	A1	A2	A3	A4	A5	A6	A7	B1	B2	B3	B4	B5	B5	B6	B-like-I	B-like-II	B-like-III	Type-4	Type-5		Type-7	Type-AmpH				
<b>Gram -</b>																										
<i>Escherichia coli</i> K12	PBP1a <i>ponA</i>	PBP1b <i>ponB</i>				PBP1c <i>pbpC</i>				PBP2 <i>pbpA</i>	PBP3 <i>ftsI</i>							PBP4 <i>dacB</i>	PBP5 <i>dacA</i>	PBP6 <i>dacC</i>	PBP6b <i>dacD</i>	PBP7 <i>pbpG</i>		PBP4b <i>yefw</i>	AmpH <i>ampH</i>	
<i>Neisseria gonorrhoeae</i> FA 1090	PBP1 <i>ponA</i>										PBP2 <i>ftsI</i>							PBP3 <i>pbp3</i>				PBP4 <i>pbp4</i>				
<b>Gram +</b>																										
<i>Bacillus subtilis</i> 168		PBP1 <i>ponA</i>	PBP2c <i>pbpF</i>	PBP4 <i>pbpD</i>		PBP2d <i>pbpG</i>		PBP3 <i>pbpC</i>	SpoVD <i>spoVD</i>	PBP2b <i>pbpB</i>	PBP2a <i>pbpA</i>	PbpH <i>pbpH</i>	PBP4b <i>yrrR</i>					PBP4a <i>dacC</i>	DacF <i>dacF</i>	PBP5 <i>dacA</i>	PBP5* <i>dacB</i>			PBP4* <i>pbpE</i>	PbpX <i>pbpX</i>	
<i>Staphylococcus aureus</i> MRSA252		PBP2 <i>pbp2</i>						PBP2a <i>mecA</i>		PBP1 <i>pbpA</i>	PBP3 <i>pbp3</i>								PBP4 <i>pbp4</i>							
<i>Listeria monocytogenes</i> 4b F236		PBP1 <i>lmo1892</i>	PBP4 <i>lmo2229</i>					PBP <i>lmo0441</i>		PBP2 <i>lmo2039</i>	PBP3 <i>lmo1438</i>								PBP5 <i>lmo2754</i>						PBP <i>lmo0540</i>	
<i>Enterococcus faecalis</i> V583		PBP1a <i>EF_1148</i>	PBP2a <i>EF_0680</i>	PBP1b <i>EF_1740</i>				PBP4 <i>EF_2476</i>		PBP2 <i>EF_0991</i>	PBP2b <i>EF_2857</i>								DacF <i>EF_3129</i>						PBP <i>EF_0746</i>	
<i>Streptococcus pneumoniae</i> R6		PBP1a <i>pbpA</i>	PBP2a <i>pbp2&amp;</i>	PBP1b <i>pbp1b</i>						PBP2x <i>pypX</i>	PBP2b <i>pbp2b</i>								PBP3 <i>pbp3</i>							

D'après (Sauvage *et al.* 2008).

Tableau 4 : Fonctions de différentes PBPs dans différentes souches.

Souche	Gène	Fonction	Technique	Type de PBP	Références
<i>Streptococcus agalactiae</i>	<i>ponA</i> (PBP1a)	Protège de la phagocytose, protéine de surface anti-phagocytaire	STM avec criblage sur des rats nouveau-nés	HMW PBP	(Jones <i>et al.</i> 2003)
<i>Escherichia coli</i> 2443	<i>mrcA</i> (PBP1a) et <i>mrcB</i> (PBP1b)	PBP1b : Formation des biofilms et motilité  PBP1a : pas de changement	Mutation ciblée (transduction P1)	HMW PBP	(Kumar <i>et al.</i> 2012)
<i>Helicobacter pylori</i>	<i>penA</i> (PBP2)	Activation des cellules dendritiques humaine et murine (temps et dose dépendant)  Induction de MHC II, CD80, CD86, p65 NF-kB  Liaison au TLR4	Production et purification de la protéine  Test in vivo et in vitro	HMW PBP	(Hill <i>et al.</i> 2011)
<i>Streptococcus thermophilus</i>	<i>pbp2b</i> (PBP2b)	Architecture du peptidoglycane, résistance contre des attaques (stress oxydatif)	Mutation aléatoire et ciblée (pGh9:ISS1, recombinaison homologue)	HMW PBP	(Thibessard <i>et al.</i> 2002)
<i>Streptococcus uberis</i>	<i>pbp2a</i> (PBP2a)	Adhérence et internalisation dans les cellules épithéliales mammaire	Mutation aléatoire et criblage (pGh9:ISS1)	HMW PBP	(Deogo <i>et al.</i> 2011)
<i>Acinetobacter baumannii</i>	<i>pbpG</i> (PBP7/8)	Virulence dans un modèle d'infection des tissus mous et un modèle de pneumonie chez le rat.	Criblage en milieu ascite de mutants obtenus par transposition.	LMW PBP	(Russo <i>et al.</i> 2009)
<i>Escherichia coli</i>	<i>dacB</i> (PBP4) et <i>pbpG</i> (PBP7)	Rôle auxiliaire (aide la PBP5) dans le maintien d'une morphologie uniforme de cellule	Mutation ciblée	LMW PBP	(Meberg <i>et al.</i> 2004)
<i>Escherichia coli</i>	<i>dacA</i> (PBP5)	Rôle principal dans le maintien du diamètre, du contour et de la morphologie de la cellule	Mutation ciblée	LMW PBP	(Nelson and Young 2000)

### III.1.4. Les PBPs de *Lactobacillus casei* ATCC 334

Le génome de *L. casei* ATCC 334 possède treize PBPs. *Lactobacillus rhamnosus* GG en possède autant, alors que les autres souches répertoriées dans le

Tableau 5 ci-dessous en possèdent moins excepté *Bacillus subtilis* TSn5 qui en possède 16. Le COG (Cluster of Orthologous Groups) et le nom des gènes des PBPs chez *L. casei* ATCC 334 sont les suivants:

COG0744 : Classe A, annotées en tant que « Membrane carboxypeptidase transpeptidase transglycosylase (penicillin-binding protein) » :

- LSEI\_0810 : ([YP\\_806088.1](#)) PBP1a.
- LSEI\_1481 : ([YP\\_806700.1](#)) PBP1a.
- LSEI\_1725 : ([YP\\_806940.1](#)) PBP1a.

COG0768 : Classe B, annotées en tant que « Cell division protein FtsI/penicillin-binding protein 2 transpeptidase » :

- LSEI\_0314 : ([YP\\_805616.1](#)) PBP2.
- LSEI\_1269 : ([YP\\_806494.1](#)) PBP2.
- LSEI\_1662 : ([YP\\_806877.1](#)) PBP2.

COG1686 : Classe C type 5

- LSEI\_0221 : D-alanyl-D-alanine carboxypeptidase ([YP\\_805526.1](#))

COG1680 : Classe C type AmpH, annotées en tant que « Beta-lactamase class C related penicillin-binding protein transpeptidase » :

- LSEI\_1152 : ([YP\\_806383.1](#)).
- LSEI\_0435 : ([YP\\_805732.1](#)).
- LSEI\_1802 : ([YP\\_807014.1](#)).
- LSEI\_1805 : ([YP\\_807017.1](#)) AmpC.
- LSEI\_2137 : ([YP\\_807329.1](#)).
- LSEI\_2773 : ([YP\\_807929.1](#)).

Tableau 5 : Comparaison du nombre de chaque COG représentant les PBPs chez différentes souches bactériennes.

	<i>Lactobacillus casei</i> ATCC 334	<i>Lactobacillus rhamnosus</i> GG	<i>Lactobacillus plantarum</i> WCFS1	<i>Lactobacillus reuteri</i> SD2112	<i>Lactobacillus johnsonii</i> NC533	<i>Lactobacillus acidophilus</i> NCFM	<i>Bacillus subtilis</i> TSn5	<i>Lactococcus lactis lactis</i> I11403	<i>Streptococcus salivarius</i> 57.1.
COG0744	3	3	2	2	3	2	4	3	3
COG0768	3	3	2	2	2	2	5	2	2
COG1680	6	6	4	2	3	6	3	1	1
COG1686	1	1	2	2	1	1	3	1	4
COG2027	0	0	0	0	0	0	1	0	0
COG4953	0	0	0	0	0	0	0	0	0
COG5009	0	0	0	0	0	0	0	0	0
<b>Total</b>	13	13	10	8	9	11	16	7	10

### III.2. Les acides téichoïques

Les acides téichoïques (TAs) représentent le second composant majeur de la paroi des lactobacilles. Ils peuvent représenter 20 à 50% du poids sec de la paroi des lactobacilles, ceci dépend de l'espèce, de la souche, du stade de croissance, du pH du milieu, de la source de carbone et de la disponibilité en phosphate... (Delcour *et al.* 1999).

Les TAs sont impliqués dans des aspects variés de la fonctionnalité de la paroi. Il existe deux types d'acides téichoïques ; ceux qui sont liés au peptidoglycane appelé WTA (wall teichoic acids) et ceux ancrés à la membrane grâce à une partie glycolipidique, il s'agit des LTA (lipoteichoic acids) (Neuhaus and Baddiley 2003). Les TAs sont des polymères anioniques composés d'une répétition d'unités de glycérol- ou de ribitol-phosphate (-Gro-P- et -Rbo-P-) constituant ainsi une source de phosphate disponible en cas de carence. Leur caractère anionique leur permet aussi de jouer un rôle en tant que piègeur de cations, surtout des ions  $Mg^{2+}$  (Hughes *et al.* 1973). Les TAs peuvent aider à créer un gradient de pH à travers la paroi en séquestrant les protons. Les LTA et les WTA sont souvent substitués avec des glycosyl- ou D-alanyl-esters. La variété des TAs est due à la nature des sucres et au nombre de résidus phosphates les composants. La composition des TA est présentée dans la Figure 24.

Les WTA sont ancrés via une unité de liaison (Gro-P)<sub>2</sub> ou <sub>3</sub>ManNAc(β1-4)GlcNAc-P qui se lie en C6 du résidu NAM du peptidoglycane. Les gènes *tag* et *tar* permettent la biosynthèse des WTA poly (Gro-P) et poly(Rbo-P) respectivement (Lazarevic *et al.* 2002). Ces gènes sont complètement absents chez *L. casei*, *Lactobacillus fermentum* et *Lactobacillus reuteri* ; ceci laisse supposer, que dans ces espèces, la présence des LTA uniquement est suffisante pour conférer toutes les fonctions biologiques importantes des TAs.

Les gènes permettant la synthèse des LTA n'ont été découverts que très récemment. En effet, l'inactivation des gènes LTA synthase *ltaS* chez *S. aureus* et *B. subtilis* permet un blocage de la biosynthèse des LTA (Grundling and Schneewind 2007; Schirner *et al.* 2009). Tous les génomes des lactobacilles contiennent au moins une copie du gène *ltaS*, une seconde copie a été identifiée chez certaines espèces, c'est le cas chez *L. casei* (LSEI\_0868 et LSEI\_1123)

(Kleerebezem *et al.* 2010). La structure des LTA de *Lactobacillus rhamnosus* GG a été étudiée, il s'avère que la chaîne des LTA comprend 50 résidus Gro-P et que les D-Alanyl esters sont les seuls substituants (74% D-Ala : Gro-P) (Perea Velez *et al.* 2007). Des LTA peuvent être retrouvés libres dans la paroi ou libérés dans le milieu extracellulaire grâce à la déacétylation du glycolipide d'ancrage. Tout comme le PG, les LTA peuvent être reconnus comme ligand par les récepteurs tels que les Toll-like receptors (TLR) présents sur les cellules épithéliales intestinales (Lebeer *et al.* 2008).

Les D-Alanyl esters des LTA et des WTA contribuent fortement à la fonction des TAs parce que leur charge positive contre partiellement la charge négative des groupements phosphates. La D-Alanylation des TAs est réalisée par un ensemble de quatre protéines codées par l'opéron *dlt* (*dltA* à *dltD*) correspondant aux gènes LSEI\_0794 à 0797. Le gène *dltA* encode pour une D-alanyl carrier protein ligase (DcI), et le gène *dltC* pour une D-alanyl carrier protein (Dcp); ces deux protéines sont impliquées dans l'activation de l'alanine par l'utilisation de l'ATP. DltB (codée par *dltB*) est une protéine transmembranaire putative impliquée dans le passage du complexe activé D-Alanyl-Dcp à travers la membrane cytoplasmique. DltD (codée par *dltD*) est une protéine membranaire facilitant la fixation de Dcp et de Dcl pour la ligation du Dcp avec D-Ala (Neuhaus and Baddiley 2003) (Figure 25). L'inactivation de gènes dans cet opéron dans de nombreuses bactéries à Gram positif a montré une diminution ou une absence complète de D-Alanyl esters sur les LTA. Ces mutants montrent une variété de changements phénotypiques tels que l'altération de la morphologie cellulaire associée à un défaut de formation du septum (Palumbo *et al.* 2006), une différence de résistance à l'autolyse (Steen *et al.* 2005), aux peptides antimicrobiens cationiques (Poyart *et al.* 2003), une modification dans la capacité d'adhésion, d'invasion de la cellule épithéliale, de virulence (Abachin *et al.* 2002), mais aussi un défaut dans la formation des biofilms (Fabretti *et al.* 2006) et une modulation de la réponse immunitaire pro- et anti-inflammatoire de l'hôte (Grangette *et al.* 2005).

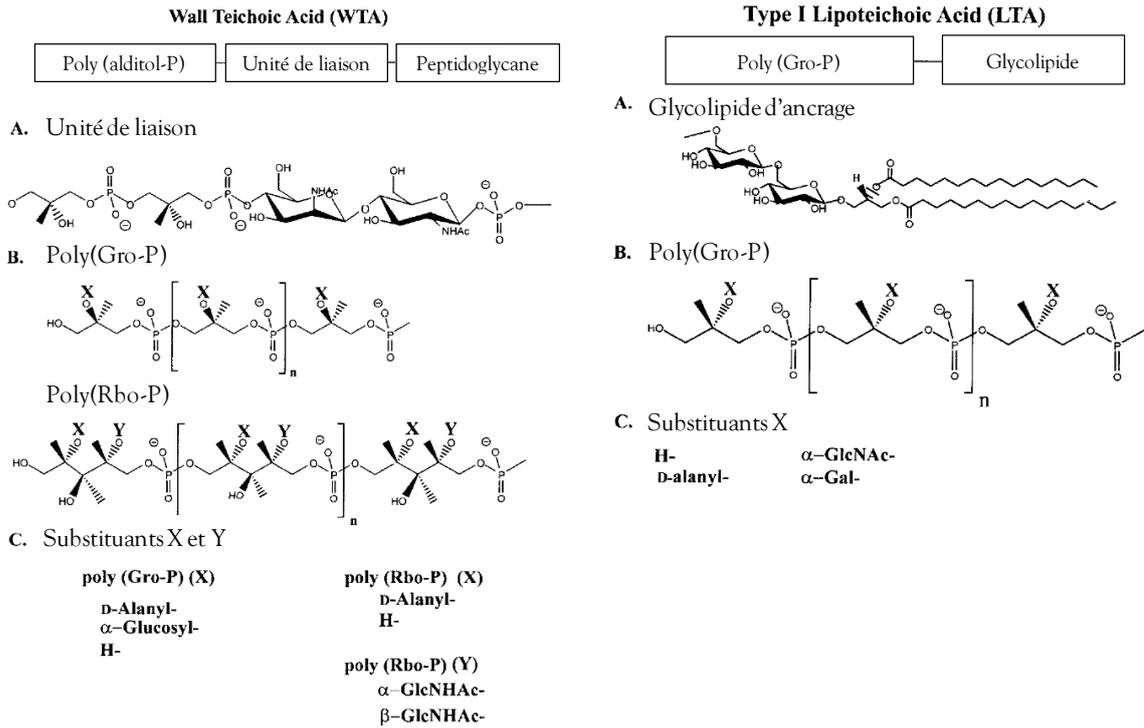


Figure 24 : Représentation de la composition des WTA et des LTA de type I.

WTA (A) Unité de liaison ; (B) Poly(Gro-P) présent chez *Bacillus subtilis* 168 et poly(Rbo-P) présent chez *S. aureus* H ; Substituents X et Y présent chez ces deux bactéries. LTA (A) Glycolipide d'ancrage ; (B) Poly(Gro-P) ; (C) Substituents X. D'après (Neuhaus and Baddiley 2003).

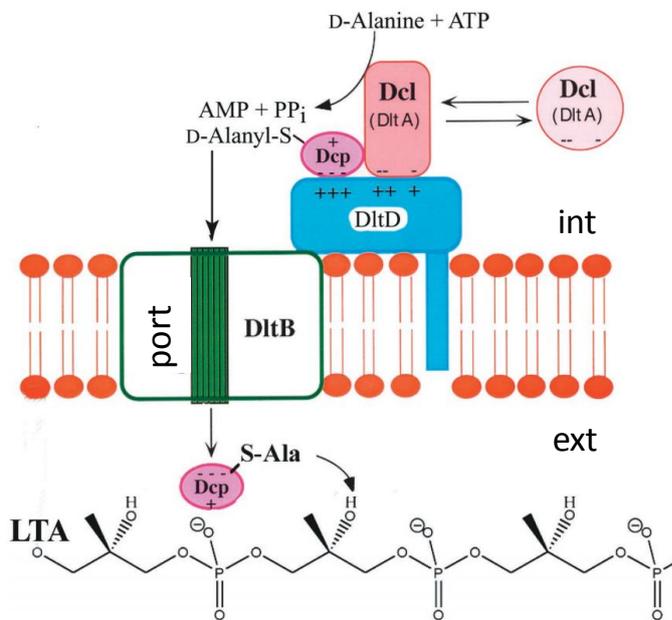


Figure 25 : Modèle d'incorporation des D-Alanyl esters sur les LTA.

DltD permet la fixation des protéines Dcp et Dcl. DltB sert de port pour la sécrétion de la D-Alanyl-Dcp à l'extérieur de la cellule où a lieu la D-Alanylation. D'après (Neuhaus and Baddiley 2003).

### III.3. Les polysaccharides de la paroi

Les polysaccharides de la paroi cellulaire sont des polysaccharides généralement neutres et peuvent être liés de manière covalente à la paroi par liaison avec un NAM. Ces polysaccharides sont alors appelés WPS (wall polysaccharides) ou ils peuvent être faiblement associés à la paroi ou libérés dans le milieu extracellulaire (appelés alors exopolysaccharides ou EPS). Lorsque les polysaccharides forment une sorte de capsule étroitement associées ou liées à l'enveloppe cellulaire, ils sont généralement nommés polysaccharides capsulaires (CPS). Cependant, la distinction entre WPS, EPS et CPS est souvent difficile et apparait quelque peu artificielle car la quantité et la localisation des polysaccharides est fortement dépendante des conditions de croissance (Kleerebezem *et al.* 2010).

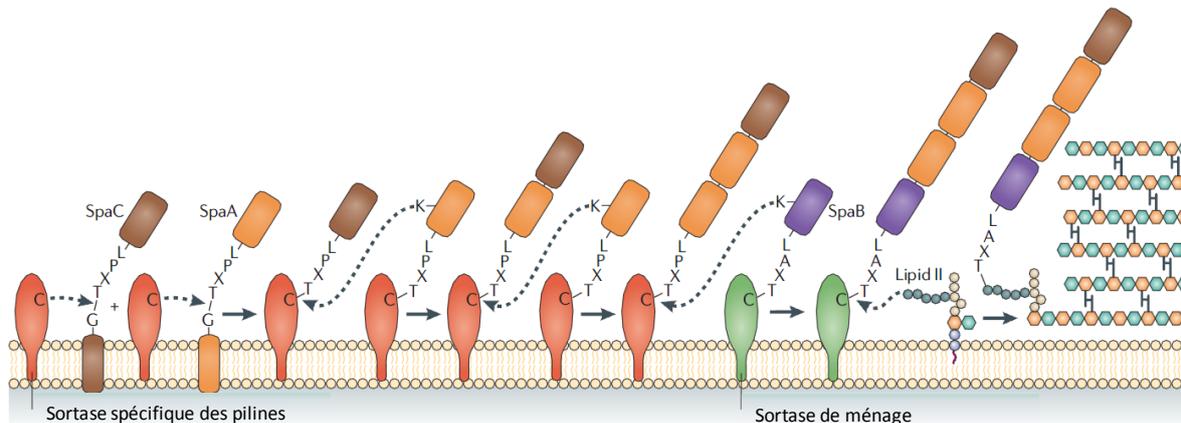
Pour les lactobacilles, le terme exopolysaccharides (EPSs) est préféré et correspond aux polysaccharides extracellulaires, qu'ils soient attachés à la paroi ou sécrétés (Lebeer *et al.* 2008). Les EPSs sont des structures complexes participant à la variété structurale de la paroi des lactobacilles. Les EPSs varient à la fois au niveau de leur composition en sucre mais aussi par leur mode de liaison, de substitution... (De Vuyst and Degeest 1999). Les EPSs des lactobacilles sont généralement des hétéropolysaccharides composés de sucres tels que le glucose, le galactose, le rhamnose, *N*-acétylglucosamine, *N*-acétylgalactosamine. D'autres résidus comme l'acide glucuronique, le glycérol-3-phosphate, des groupements phosphate, acétyl ou pyruvyl peuvent être présents. Certains lactobacilles comme *Lactobacillus reuteri* et *Lactobacillus acidophilus* sont capables de synthétiser des homopolysaccharides tels que, respectivement, les glucanes et les fructanes (Tieking *et al.* 2005).

Les gènes codant pour la biosynthèse des EPSs sont généralement regroupés dans un cluster de gène de 12 à 25 kb. Chez *L. rhamnosus*, le cluster mesure 18,5 kb et comporte 17 gènes (Peant *et al.* 2005). Deux classes d'EPS ont été identifiées chez cette espèce ; un type EPS long et riche en galactose et un type d'EPS court et riche en glucose/mannose (Francius *et al.* 2008). La contribution des EPSs dans la fonctionnalité de la paroi cellulaire n'est pas claire, cependant leur rôle général est de participer à l'interaction et la protection des lactobacilles avec les composants environnementaux et de permettre l'adhésion et la formation de biofilms sur des surfaces biotiques et abiotiques (Lebeer *et al.* 2011).

### III.4. Les pili et flagelles

Les pili sont des protéines polymériques (hétérotrimères) composées, en général, d'une piline majeure correspondant en quelque sorte à la base du pilus, d'une piline mineure décorant la piline majeure et d'une autre piline mineure comportant des propriétés adhésives et se situant souvent à l'extrémité du pilus. Les pili des Gram positifs ont un diamètre assez fin (1 à 10 nm) mais peuvent cependant atteindre une longueur de 1 µm, voire plus. Chaque piline est

constituée de sous-unités de piline. Ces sous-unités sont liées de manière covalente l'une à l'autre grâce à l'activité transpeptidase de sortase spécifique à la piline. Ces liaisons sont formées entre la thréonine d'un motif LPxTG et la lysine d'un motif YPKN de piline. Une autre sortase (sortase de ménage) nommée en général StrA permet la liaison des pili au peptidoglycane en reconnaissant le motif LPXTG. (Kankainen *et al.* 2009; Mandlik *et al.* 2008) (Figure 26).



**Figure 26 : Représentation du mécanisme d'assemblage des pili hétérotrimériques.**

Une sortase spécifique des pilines clive le peptide signal (motif LPxTG) des pilines SpaC et SpaA au niveau de la thréonine générant ainsi des acyl-enzyme intermédiaire. Un pont peptidique est ensuite formé entre les motifs LPxTG et YPKN de SpaC et SpaA. A ce dimère SpaC-SpaA sont ensuite ajoutés des sous-unités SpaA additionnelles formant ainsi la polymérisation du pilus. Le pilus est ensuite ancré à la paroi grâce à la sous-unité SpaB et une sortase de ménage en passant par un intermédiaire acyl-lipide II.

Les gènes codant pour les trois sous-unités de piline et pour la sortase sont généralement localisés dans le même locus. Des éléments transposables sont fréquemment retrouvés aux extrémités de ce locus, ce qui laisse suggérer une origine par un transfert horizontal de gène. Deux loci (*spaCBA*, *spaFED* et *strC1* et *strC2* codant pour leur sortase respective) ont été identifiés chez *L. rhamnosus* GG comme intervenant dans la synthèse des trois sous-unités de piline et de leur sortase. *spaA* et *spaD* correspondent aux sous-unités majeures formant la base du pilus, *spaB* et *spaE* aux sous-unités mineures auxiliaires, *spaC* et *spaF* aux sous-unités mineures de grande taille avec une activité d'adhésine et *srtC1* et *srtC2* aux sortases. Les sous-unités SpaB, SpaC et SpaF jouent un rôle dans l'adhésion au mucus intestinal (Kankainen *et al.* 2009; von Ossowski *et al.* 2010). Chez *L. casei* ATCC 334, les loci LSEI\_0455 à 0457 et LSEI\_2362 à 2366 correspondraient respectivement aux loci *spaCBA* et *spaFED* de *L. rhamnosus*. Les pili ont un rôle dans l'adhésion, l'invasion, l'agrégation, la formation de biofilm et l'immunomodulation (Danne and Dramsi 2012; Lebeer *et al.* 2012).

La présence de flagelles est une caractéristique inhabituelle chez les lactobacilles mais existe chez 12 espèces (Neville *et al.* 2012). La majeure partie (11 espèces sur 12) a été isolée d'aliment ou de l'environnement. Les flagelles sont des structures protéiques longues et fines qui forment des hélices rigides qui tournent dans le sens inverse des aiguilles d'une montre

afin de permettre le déplacement de la bactérie. Ils sont composés de polymères de flagelline qui sont reconnus par les TLR5 (Toll-like receptor 5) de l'hôte et vont activer des voies de signalisation et moduler la réponse immunitaire de l'hôte (Tallant *et al.* 2004).

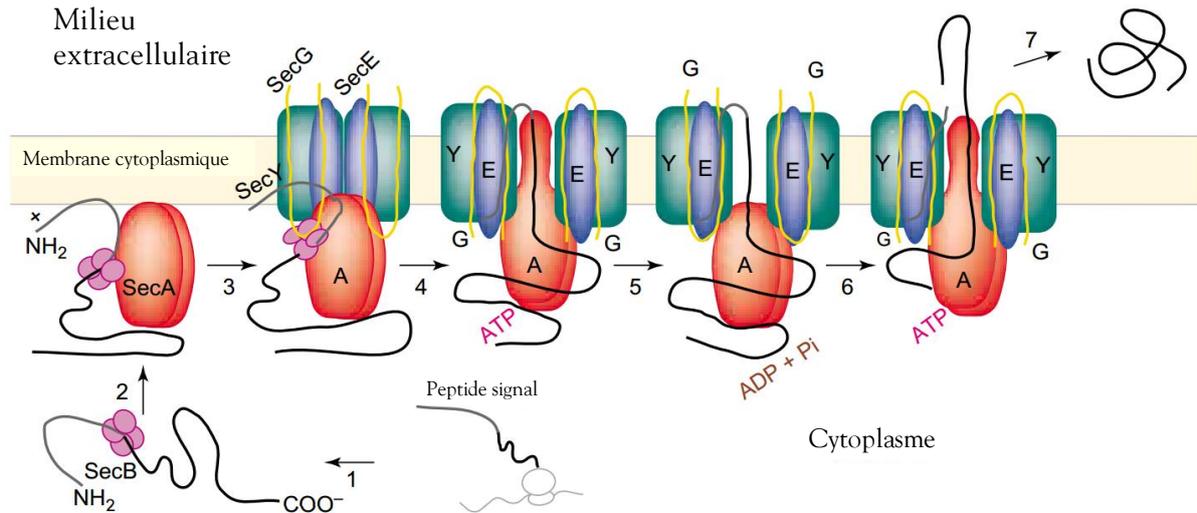
### **III.5. Les protéines de surface**

L'exoprotéome des lactobacilles contient deux grands groupes de protéines : les protéines sécrétées et les protéines associées à la surface cellulaire. Ce dernier groupe peut être divisé en plusieurs sous-catégories :

- Les protéines ancrées dans la membrane cytoplasmique grâce à un domaine hydrophobe C- ou N-terminal ;
- Les lipoprotéines ancrées dans la membrane grâce à leur domaine lipidique N-terminal ;
- Les protéines liées de manière covalente au peptidoglycane grâce à un motif LPxTG en C-terminal ;
- Les protéines liées de manière non-covalente à la surface cellulaire grâce à différents domaines de liaison ou attachées à une autre protéine de la paroi par l'intermédiaire d'interaction protéine-protéine.

#### **III.5.1. Les protéines ancrées dans la membrane cytoplasmique**

La voie majeure de sécrétion chez les lactobacilles est appelée Sec ; elle permet le passage des protéines à travers la membrane cytoplasmique (Driessen and Nouwen 2008). Cette translocation est possible grâce au pore formé par les protéines SecYEG et grâce à la protéine motrice ATPase (SecA) (Figure 27). Cet ensemble appelé translocase est associé à un complexe hétérotrimérique (SecDF-YajC) permettant la régulation de l'activité de SecA. Ce complexe permet aussi de lier la protéine YidC nécessaire à l'insertion membranaire. La protéine SecB est une chaperonne dédiée à l'export de protéine.



**Figure 27 : Représentation du processus de translocation de protéines.**

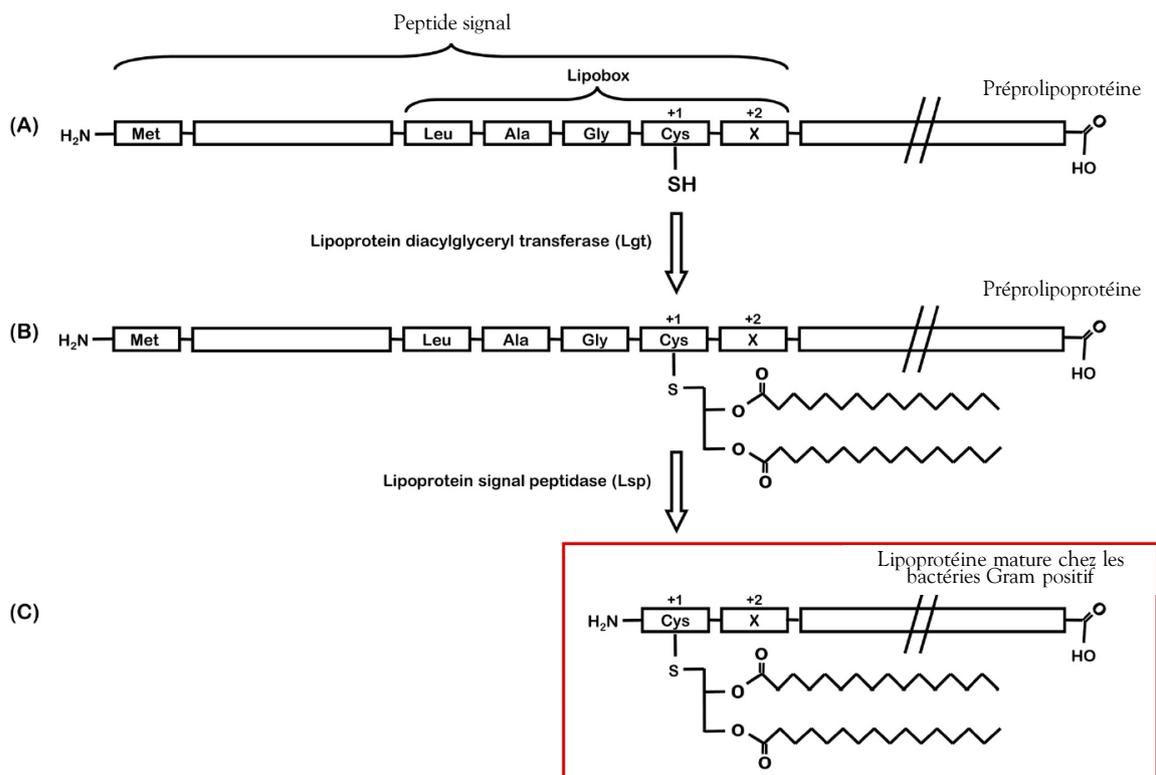
La préprotéine est représentée par une ligne noire et le peptide signal par une région grise. Etapes 1 à 3 : le peptide signal est reconnu par la machinerie Sec. La protéine chaperonne SecB permet un bon repliement de la protéine pour qu'elle soit reconnue par SecA. La protéine SecA portant la pré-protéine se lie à la membrane. Les protéines SecY, G et E forment le port. Etapes 4 et 5 : L'étape d'initiation nécessite la présence d'ATP mais pas son hydrolyse. Etape 6 : La translocation requiert des cycles d'hydrolyse d'ATP et/ou une force proton-motrice. Etape 7 : Libération et/ou repliement de la protéine dans l'espace périplasmique. D'après (Mori and Ito 2001).

Les protéines utilisant ce système de sécrétion possèdent un peptide signal N-terminal composé de 3 régions : (1) la région N qui est chargée positivement en N-terminal ; (2) la région H composée de 15 à 25 résidus hydrophobes et enfin (3) la région C qui contient le site reconnu par les peptidases signal (SPases). Deux types de SPases existent, SPases de type I reconnaissent le site de clivage AxA (van Roosmalen *et al.* 2004) et les SPases de type II reconnaissent le site de clivage [LVI][ASTVI][GAS]C appelé lipobox présent dans les lipoprotéines (Kovacs-Simon *et al.* 2011). Le peptide signal est reconnu par les SPases et clivé, permettant ainsi la sécrétion de la protéine. Cependant, la plupart des protéines ne possède pas ce peptide signal ou possède un motif similaire qui n'est pas clivé, les protéines restent alors ancrées dans la membrane cellulaire (Zhou *et al.* 2008). Ces protéines sont principalement impliquées dans des processus extracellulaires tels que le transport, le métabolisme de la paroi, la compétence, la transduction du signal et le renouvellement protéique (Kleerebezem *et al.* 2010).

Dans certains cas, même si les protéines possèdent un site de clivage conforme, elles peuvent rester ancrées à la membrane grâce à un domaine C-terminal transmembranaire. Les génomes des lactobacilles codent pour un nombre variable de ces protéines, dont beaucoup n'ont pas de fonction connue (Kleerebezem *et al.* 2010).

### III.5.2. Les lipoprotéines

La synthèse des lipoprotéines est réalisée à partir d'un précurseur appelé préprolipoprotéine qui possède un peptide signal N-terminal composé de la lipobox. Lors de la maturation des lipoprotéines chez les bactéries Gram positif, un diacylglycéryl est ajouté sur le groupement thiol de la cystéine présent dans la lipobox grâce à la lipoprotein diacylglycéryl transferase (Lgt). Ce diacylglycéryl va permettre l'ancrage dans la membrane cytoplasmique. Après lipodation, la Spase de type II ou lipoprotein signal peptidase (Lsp) va cliver le peptide signal. Le résidu cystéine devient alors le nouveau résidu N-terminal, cet ensemble forme ainsi la lipoprotéine mature (Kovacs-Simon *et al.* 2011) (Figure 28). Les lipoprotéines encodées par les génomes des lactobacilles sont principalement des protéines de liaison au substrat des ABC-transporteurs, mais sont aussi impliquées dans l'adhérence, la résistance aux antibiotiques, l'homéostasie de l'enveloppe cellulaire, la sécrétion, le repliement et la translocation (Kleerebezem *et al.* 2010). *L. casei* ATCC 334 possède 46 lipoprotéines d'après l'analyse par le logiciel LocateP (<http://www.cmbi.ru.nl/locatep-db/cgi-bin/locatepdb.py>).



**Figure 28 : Représentation schématique de la synthèse des lipoprotéines.**

(A) Préprolipoprotéine ; (B) Lipodation de la préprolipoprotéine en diacylglycéryl préprolipoprotéine ; (C) Clivage du peptide signal et obtention de la lipoprotéine mature. D'après (Kovacs-Simon *et al.* 2011).

### III.5.3. Les protéines ancrées à la paroi avec un domaine LPxTG.

Les protéines avec motif LPxTG possèdent une séquence en C-terminale composée des résidus Leucine (L), Proline (P), x (représentant n'importe quel acide aminé), Thréonine (T), et Glycine (G). Cette séquence peut être sujette à certaines variations d'acides aminés suivant l'espèce bactérienne (Boekhorst *et al.* 2005). Ces protéines sont ancrées à la paroi grâce à cette séquence qui est reconnue par la sortase de ménage (StrA). Ces protéines ne doivent pas uniquement posséder ce motif LPxTG pour être ancrées au peptidoglycane ; en effet, elles doivent aussi contenir un peptide signal en région N-terminale et un domaine hydrophobe situé juste après le motif LPxTG ainsi qu'une queue composée en majorité de résidus chargés positivement (Schneewind *et al.* 1992) (Figure 29A). Le peptide signal permet la sécrétion de la protéine par la voie Sec alors que la partie C-terminale chargée positivement va permettre l'ancrage dans la membrane cellulaire. Une fois la protéine ancrée dans la membrane, la sortase StrA va réaliser la réaction de transpeptidation requise pour l'ancrage au peptidoglycane. La première étape de la réaction de transpeptidation est le clivage de la protéine entre les résidus thréonine et glycine du motif LPxTG. De ce clivage va résulter un complexe composé de la sortase et de la protéine. Par la suite, la liaison acyl-thioester entre ces deux protéines va être soumise à une attaque nucléophile permettant ainsi le lien entre la protéine et le pentapeptide du lipide II (précurseur du peptidoglycane). La protéine est alors insérée dans le peptidoglycane lors de la synthèse de ce dernier (Call and Klaenhammer 2013) (Figure 29B). *L. casei* ATCC 334 possède 21 protéines avec un motif LPxTG d'après l'analyse locateP, les différentes variations du motif chez *L. casei* sont présentées dans la Figure 30. Une des protéines (codé par le gène LSEI\_2431) contient, en plus du domaine LPxTG, deux domaines de liaison au collagène (collagen binding domain et Cna protein B-type domain respectivement PF05737 et PF05738).

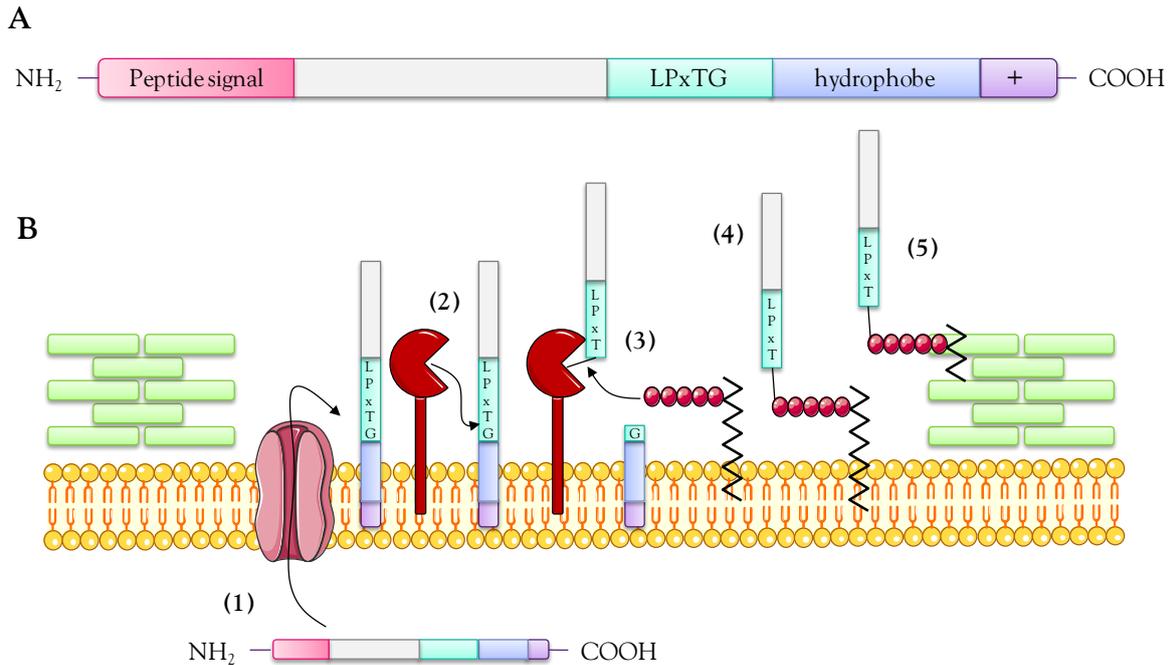


Figure 29 : Ancre des protéines à motif LPxTG.

(A) Protéine reconnue comme substrat par les sortases, présence d'un peptide signal en N-terminal, d'un motif LPxTG en C-terminal ainsi qu'une partie hydrophobe et d'un domaine composé de résidus chargés positivement. (B) Transfert des protéines LPxTG du cytoplasme à l'ancrage sur le peptidoglycane : (1) le système de sécrétion représenté en rose reconnaît le peptide signal présent sur la protéine et la transfère du cytoplasme vers le milieu extracellulaire, ancrage de la protéine dans la membrane grâce à son domaine hydrophobe et ses résidus chargés positivement. (2) La sortase de ménage reconnaît le motif LPxTG et clive la protéine entre T et G. (3) Le complexe sortase/protéine est attaqué par le pentapeptide du lipide II. (4) Création du complexe protéine-lipide II. (5) Synthèse du peptidoglycane et ancrage. D'après (Call and Klaenhammer 2013).

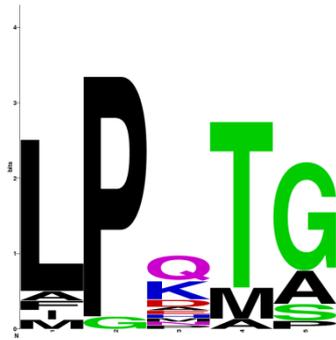


Figure 30 : Logo représentant les variations du motif LPxTG chez *L. casei* ATCC 334.

### III.5.4. Les protéines liées de manière non-covalente à la paroi

Certaines protéines peuvent être liées de façon non covalente à la paroi grâce à différents domaines dont les majoritaires sont détaillés ci-dessous.

#### Les protéines à domaine WxL

Le **domaine WxL** (pfam13731) a été à l'origine prédit grâce aux analyses *in silico* des génomes de différents lactobacilles tels que *L. plantarum* (Kleerebezem *et al.* 2003), *L. coryniformis* (Schachtsiek *et al.* 2004), et *L. sakei* (Chaillou *et al.* 2005). Ce domaine

caractérisé chez *Enterococcus faecalis* est composé de deux motifs, un motif hautement conservé Yxxx(L/I/V)TWxLxxxP situé dans les 130 derniers résidus C-terminal des protéines à domaine WxL et un second motif en amont correspondant aux résidus WxL (Brinster *et al.* 2007). Ce domaine est nécessaire pour la liaison non covalente avec la paroi. Les gènes codant pour ces protéines possédant ce domaine ont été retrouvés dans des clusters de gènes codant pour d'autres protéines extracellulaires (motif LPxTG) laissant supposer que cet ensemble de protéines formerait un complexe extracellulaire impliqué dans l'utilisation des glucides (Siezen *et al.* 2006). *L. casei* ATCC 334 possède 9 protéines avec un domaine WxL.

#### Les protéines à domaine LysM

Le **domaine LysM** (motif lysine) (pfam01476) par exemple, est retrouvé dans de nombreuses enzymes extracellulaires qui sont impliquées dans le métabolisme de la paroi telles que les peptidoglycan hydrolases (PGH) (Buist *et al.* 2008). Le domaine LysM peut être présent en N- ou C-terminal mais également dans la partie centrale de la protéine. La taille de ce motif peut aller de 44 à 65 résidus d'acides aminés, les 10 voire 16 premiers résidus sont assez bien conservés. La région centrale est cependant, elle, peu conservée exceptées l'isoleucine/leucine en position 23 et 30 et l'asparagine en position 27. Certaines protéines peuvent posséder plusieurs domaines LysM. Le point isoélectrique de ce domaine peut varier de 4 à 12 (Buist *et al.* 2008). *L. casei* ATCC 334 possède quatre protéines avec un domaine LysM dont une qui en possède deux.

#### Les protéines à domaine SH3

Le **domaine SH3**, nommé SH3b (pfam08239) chez les procaryotes semble être un domaine de liaison à la paroi. Les protéines étudiées possédant ce domaine sont généralement des protéines impliquées dans la lyse du peptidoglycane (Gu *et al.* 2011; Navarre *et al.* 1999). *L. casei* ATCC 334 possède trois protéines avec un domaine SH3 : la protéine Lysozyme M1 (1,4-beta-N-acetylmuramidase) codée par le gène LSEI\_0599 qui porte le domaine SH3\_5, la protéine N-acetylmuramoyl-L-alanine amidase codée par le gène LSEI\_1536 qui porte le domaine SH3\_3 présent en trois copies et la protéine 1,4-beta-N-acetylmuramidase codée par le gène LSEI\_0236 qui porte le domaine SH3\_8 présent en sept copies.

#### Les protéines à domaine putatif de liaison au peptidoglycane

Le **domaine putatif de liaison au peptidoglycane** (pfam 01471) est composé de trois hélices  $\alpha$  et est retrouvé en région N ou C-terminale d'une variété d'enzymes impliquées dans la dégradation de la paroi cellulaire. *L. casei* ATCC 334 possède une seule protéine avec un domaine putatif de liaison au peptidoglycane.

#### Les protéines de la couche S

La **couche S** est constituée de nombreuses sous-unités (glyco)protéiques identiques. Ces sous-unités souvent riches en acides aminés acides et hydrophobes sont liées entre elles et attachées

à la surface par des liaisons non-covalentes. Ces protéines possèdent un peptide signal de 25 à 32 résidus d'acides aminés permettant leur sécrétion par la voie générale. La séquence en acides aminés de ces protéines matures varie cependant considérablement ; une analyse par microscopie électronique est alors nécessaire pour déterminer la présence d'une couche S à la surface de la bactérie (Hynonen and Palva 2013; Sara and Sleytr 2000). La base de données Pfam contient néanmoins différents domaines responsables de l'ancrage non-covalent à la paroi tels que SLH (pfam00395), SLAP (pfam03217), S\_layer\_N (pfam05123) et S\_layer\_C (pfam05124) (Kleerebezem *et al.* 2010). La fonction exacte de la couche S n'est pas encore complètement comprise mais il semblerait qu'elle protège la bactérie des environnements hostiles et aide à maintenir l'intégrité cellulaire (Avall-Jaaskelainen and Palva 2005), à adhérer aux cellules épithéliales de l'hôte (Buck *et al.* 2005) et à inhiber l'adhésion de bactéries pathogènes (Zhang *et al.* 2010). Aucune protéine possédant un des domaines responsables de l'ancrage n'a été retrouvée chez *L. casei* ATCC 334.

#### Les protéines à domaine de liaison à la choline

Le **domaine de liaison à la choline** (pfam01473) est composé de 2 à 10 répétitions de 20 résidus d'acides aminés comprenant des résidus aromatiques et des glycines. Les protéines possédant ce domaine sont capables de se lier à la choline des phosphoryl-cholines des LTAs et des WTAs (Hakenbeck *et al.* 2009). Ce domaine est souvent retrouvé dans la séquence d'enzymes extracellulaires telles que les autolysines et les muramidases (Kleerebezem *et al.* 2010). Aucune protéine avec le domaine de liaison à la choline n'a été retrouvée chez *L. casei* ATCC 334.

# **Contexte, Objectifs et Stratégies**

Ce travail de thèse s'inscrit dans le cadre du contrat européen ERC « HOMEOPITH » : « Homeostasis and rupture of the gut epithelium in the presence of commensals and pathogens » coordonné par Philippe Sansonetti. Un des axes de recherche de ce projet était de déterminer comment les commensaux et les pathogènes pouvait affecter la signalisation immunitaire innée conduisant ainsi à la tolérance de la flore résidente et au rejet inflammatoire des pathogènes. Le travail de recherche concernant la bactérie commensale *Lactobacillus casei* a commencé au sein de l'UMR PAM et en collaboration avec l'équipe PMM de l'Institut Pasteur en 2009 avec le Post-Doctorat d'Hélène Licandro-Seraut. L'objectif était alors d'identifier les gènes intervenant dans les mécanismes de commensalisme chez *L. casei* dans le tube digestif des mammifères par une démarche globale de génétique inverse, mettant notamment en œuvre la mutagenèse aléatoire par transposition et l'approche STM (Signature Tagged Mutagenesis) (Mazurkiewicz *et al.* 2006) afin de réaliser le criblage des mutants *in vivo* (c'est-à-dire chez un modèle animal en vue d'identifier les mutants incapables de s'établir dans l'intestin). La STM développée pour les pathogènes (West *et al.* 2005) permet d'inoculer simultanément un ensemble de plusieurs mutants puis de tracer chacun des mutants dans un environnement complexe grâce à différents tags (séquences de 40 pb) présent sur le transposon intégré au génome des mutants.

*L. casei* est un bacille Gram-positif, anaérobie facultatif, non mobile et non sporulant. Cette bactérie lactique, hétérofermentaire facultative, a été isolée dans de nombreux environnement tels que les produits laitiers comme les laits fermentés et le fromage, les végétaux comme les olives vertes fermentés mais aussi dans le tractus intestinal et l'appareil reproducteur humains et d'animaux (Kandler and Weiss 1986). *L. casei* ATCC 334 (CIP 107868) est depuis 1996 (Dicks *et al.* 1996) proposé comme le néotype des *L. casei*. De plus, son génome a été séquencé (GenBank: NC\_008526.1 et NC\_008502.1), il a donc été choisi comme modèle pour ce projet.

L'équipe a tout d'abord testé plusieurs vecteurs de mutagenèse par transposition mais ces vecteurs se sont avérés inefficaces chez *L. casei* malgré une optimisation des conditions expérimentales de transformation de la souche. Par la suite, lors d'une collaboration avec Pascale Serror (INRA de Jouy en Josas), un système de transposition utilisant deux vecteurs complémentaires fonctionnant en trans et opérationnels chez *L. casei* a été mis au point. J'ai pu, lors de mon stage de Master Recherche, contribuer à la validation du caractère aléatoire de la transposition du vecteur sur le génome de *L. casei*. Une partie de l'objectif de l'axe étant réalisé, il fallait alors identifier les gènes intervenant dans les mécanismes de commensalisme chez *L. casei* dans le tube digestif. Les objectifs de ma thèse ont donc été multiples (Figure 31) :

Le premier objectif était de rendre la banque de mutants STM de *L. casei* doublement informative, c'est-à-dire de connaître les gènes impliqués dans un phénotype mais aussi les gènes qui ne le sont pas. Pour cela, le site d'insertion du transposon devait être identifié pour

tous les mutants testés. Nous devions pour atteindre cet objectif, développer une méthode d'extraction d'ADN à grande échelle et de séquençage adaptées.

Le second objectif était de réaliser le criblage fonctionnel de la banque de mutants de *L. casei* pour leur capacité (ou incapacité) à s'établir dans l'intestin. Ce travail devait être réalisé pour moitié par Hélène Licandro-Seraut et l'autre moitié par moi. Nous devions (i) réorganiser la banque de mutants STM en pool en fonction du tag de chaque mutant mais aussi en fonction du site d'insertion du transposon (région génique ou intergénique), (ii) réaliser la compétition des mutants dans l'intestin *in vivo*, (iii) quantifier les mutants par qPCR, avant et après compétition *in vivo*, grâce au tag présent dans le transposon, (iv) identifier les mutants ayant un phénotype d'établissement dans l'intestin différent et par conséquent les gènes impliqués dans le phénotype grâce au séquençage de la banque.

Enfin, le dernier objectif de la thèse était de caractériser fonctionnellement une ou plusieurs protéines clés impliquées dans l'établissement dans l'intestin. Au vu des résultats obtenus suite à la réalisation du second objectif, nous avons choisi de nous intéresser plus particulièrement aux trois mutants d'un opéron composé d'un système à deux composants et d'une protéine impliquée dans la biosynthèse du peptidoglycane et de déterminer pourquoi ces mutants ne pouvaient s'établir dans l'intestin. En ce qui concerne le système à deux composants, nous voulions vérifier sa fonctionnalité (phosphorylation et transfert de phosphate), identifier le motif ADN de reconnaissance du RR et déterminer les conditions d'induction du système. Pour la protéine impliquée dans la biosynthèse du peptidoglycane, nous voulions vérifier sa fonctionnalité (activité enzymatique) et déterminer l'impact de l'absence de cette protéine sur la bactérie et plus particulièrement sur la paroi cellulaire.

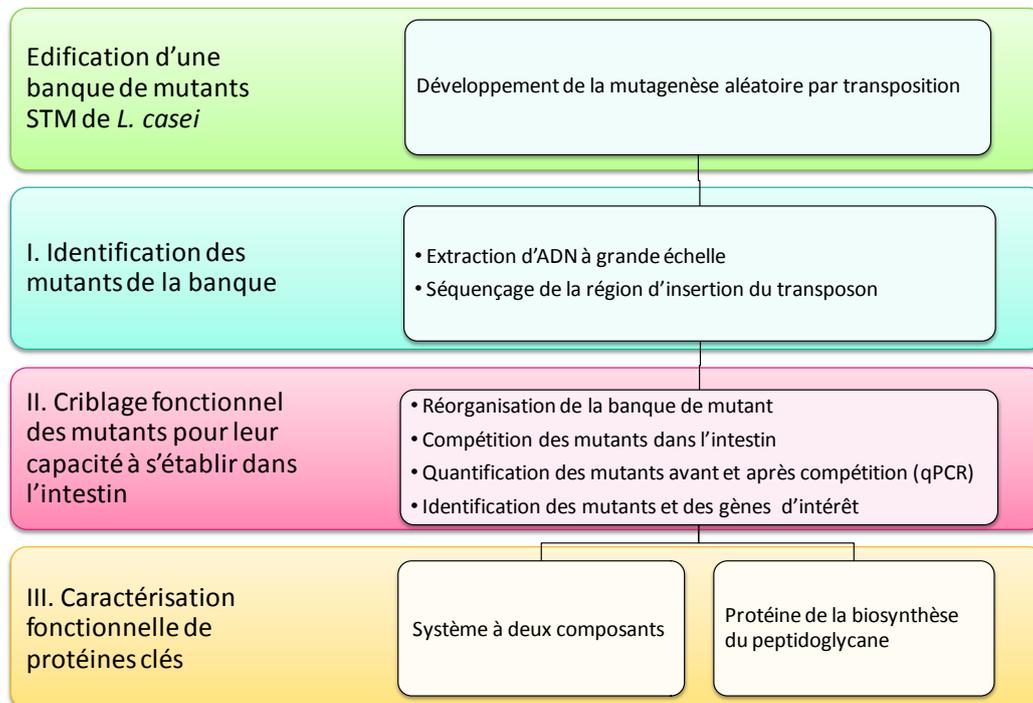


Figure 31 : Schéma des principaux objectifs du travail de thèse et des stratégies mises en œuvre.

# **Matériels et Méthodes**

## I. Souches bactériennes et conditions de culture

*Lactobacillus casei* ATCC 334 (CIP 107868), fournie par la Collection de l'Institut Pasteur, a été cultivée à 37°C dans du milieu MRS liquide (De Man *et al.* 1960) sans agitation ou sur milieu MRS gélosé. *Escherichia coli* TG1 et BL21 (DE3) ont été cultivées à 30 ou 37°C sur milieu LB gélosé (Bertani 1951) ou LB liquide avec agitation en aérobiose. Le détail des souches et leurs caractéristiques sont présentés dans le Tableau 8. Les antibiotiques et leurs concentrations utilisées pour les mutants de ces deux espèces sont donnés dans le Tableau 6.

Tableau 6 : Concentration des antibiotiques utilisés

Souches	Antibiotiques (mg.L <sup>-1</sup> )			
	Amp	Cm	Erm	Kan
<i>E. coli</i>	200	20	200	50
<i>L. casei</i>	-	5	5	-

Amp : Ampicilline ; Cm : Chloramphénicol ; Erm : Erythromycine ; Kan : Kanamycine

## II. Techniques générales de manipulations des acides nucléiques

L'extraction, la préparation et la manipulation générale des acides nucléiques sont réalisées selon les méthodes décrites par Sambrook *et al.* (Sambrook *et al.* 1989). Les techniques sont décrites brièvement ci-après.

### II.1. Extraction des acides nucléiques

#### II.1.1. Extraction d'ADN total

Cinq millilitres de culture (cultivée pendant 16 h) ont été centrifugés à 3220 x g. Les cellules obtenues ont été lysées par du lysozyme (10 mg.ml<sup>-1</sup>) à 37°C pendant 1 h pour *L. casei*. Puis 50 µl de SDS 10% (m/v) ont été ajoutés pour déstructurer les membranes. Les protéines et les déchets ont été précipités par du NaClO<sub>4</sub> 5 M, puis éliminés par extraction au phénol-chloroforme ou au chloroforme pur. L'ADN total a été précipité par l'éthanol 95% (v/v) à -20°C, lavé 2 fois par l'éthanol 70% (v/v), séché à l'air (environ 10 min), puis repris dans du tampon TE (Tris-HCl 10 mM, EDTA 1mM, pH 7,5) et congelé en aliquotes à -20°C jusqu'à utilisation.

Les extractions d'ADN total de *L. casei* peuvent aussi être réalisées avec le kit NucleoSpin® Tissue (Macherey-Nagel), grâce à une mise au point du protocole présentée dans l'article 2 des résultats « Rapid 96-well plates DNA extraction and sequencing procedures to identify genome-wide transposon insertion sites in a difficult to lyse bacterium: *Lactobacillus casei*. ».

### II.1.2. Extraction d'ADN plasmidique

Les extractions d'ADN plasmidique ont été réalisées avec le kit EZ-10 Spin Column Plasmid DNA Minipreps (BioBasic Inc) selon les recommandations du fournisseur. Cette extraction est basée sur la lyse alcaline des cellules bactériennes avec dénaturation et renaturation de l'ADN plasmidique, suivi de l'adsorption de celui-ci sur la membrane silice en présence de haute teneur en sel. Les colonies présentes sur un quart de boîte de pétri après incubation pendant 12 h ont été utilisées. L'ADN plasmidique a enfin été élué dans 50 µl de tampon TE (faible teneur en sel), ou dans 50 µl d'eau MilliQ en cas d'envoi pour séquençage par GATC Biotech.

### II.1.3. Extraction des ARN totaux

Le nombre des cellules nécessaires (la quantité optimale étant l'équivalent de 50 ml de culture à DO<sub>600nm</sub> de 0,6) en milieu de condition adéquate (phase de croissance et condition testée : pH, Bile, NaCl, ...) est obtenu par centrifugation puis resuspension dans 1 ml de TriReagent (Sigma). Les ARN totaux sont extraits par casse aux billes (billes de verre de 50 µm de diamètre dans une proportion de 1 volume de billes pour 4 volumes de suspension cellulaire) dans un FastPrep®-24 (MP) quatre fois à puissance 6,5 pendant 40 secondes. Une alternance de 1 min dans l'eau glacée entre chaque passage est obligatoire pour éviter la dégradation des ARNs par échauffement. Ensuite, la purification des ARN totaux a été réalisée selon les recommandations du fabricant. Le broyat cellulaire dans le TriReagent a été extrait par 1/5 du volume de chloroforme pur et la phase aqueuse a été récupérée. Les ARNs ont été précipités par un volume d'isopropanol à température ambiante et récoltés par centrifugation. Le culot a été lavé deux fois dans de l'éthanol à 70% (v/v), séché et repris dans 50 µl d'eau DEPC certifiée sans DNase et RNase (Invitrogen™).

## II.2. Amplification de l'ADN par PCR

La PCR (Polymerase Chain Reaction) a été utilisée pour amplifier les fragments d'ADN spécifiques. De façon générale, les réactions de PCR ont été réalisées dans un thermocycleur Bio-Rad C1000™. Le mélange réactionnel contenait généralement : la matrice d'ADN (environ 0,1 µg), 0,02 unités de Phusion High-Fidelity DNA polymerase (Thermo Scientific), 100 µM de chaque désoxyribonucléotide triphosphate (dATP, dCTP, dGTP, dTTP), 0,5 µM

de chaque oligonucléotide. Les oligonucléotides (Tableaux 7, 3 de l'article 4 et S3 de l'article 3) utilisés durant cette thèse ont été conçus à l'aide du logiciel Oligo-Analyzer 1.0.3 et du logiciel Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>) et synthétisés par Eurogentec. La programmation des cycles de température a été généralement la suivante : après une première dénaturation de l'ADN pendant 2 min à 95°C, 30 cycles de dénaturation de l'ADN double brin (98°C, 10 sec), hybridation des amorces avec l'ADN (55 à 60 °C selon les amorces, 20 sec) et synthèse de l'ADN (72°C, 20 sec à 2 min selon la longueur de la séquence amplifiée). Les produits de PCR ont été analysés par électrophorèse sur gel d'agarose de 1 à 2% (m/v). La taille des fragments d'ADN amplifiés (amplicons) a été estimée par étalonnage du gel de 200 bp à 10.000 bp avec marqueur SmartLadder (Eurogentec). Les produits de PCR ont été purifiés sur des mini colonnes de purification (EZ-10 Spin Column PCR Product Purification kit, BioBasic Inc).

**Tableau 7 : Amorces utilisées pour la construction des plasmides de surexpression des systèmes à deux composants.**

Nom	Séquence 5'→3'	Enzyme de restriction	Taille du produit PCR
Surexpression avec His-tag N-terminal			
Primer0219F	GCCTGGTGCCGCGCGGCAGC <b>CATATG</b> AAAAATTTAATTGTTGACGA	NdeI	736
Primer0219R	CAGTGGTGGTGGTGGTGGT <b>CTCGAG</b> TTATTAGGCCTCAACCTTATAGC	XhoI	
Surexpression avec His-tag C-terminal			
LSEI_0220FHAMPNcoI	ATA <b>CCATGG</b> GCACCGTCAATTCAATGGCC	NcoI	
LSEI_0220RXhoI	GTG <b>CTCGAG</b> TTTTGACTTTCCTGCTTCCT	XhoI	1205
LSEI_0219FbisNcoI	ATA <b>CCATGG</b> GCAAAAATTTAATTGTTGA	NcoI	
LSEI_0219RXhoI	GTG <b>CTCGAG</b> GGCCTCAACCTTATAGCCGA	XhoI	701
pET28aPCRR	AGATTGGTATCGCTGCTCTCTG <b>CCATGG</b> TATATCTCCTTCTTAAAGTTAAAC	NcoI	
pET28aPCRf	GCGATTCTTCAGCAATATTT <b>CTCGAG</b> CACCACCACCACCACCTGAG	XhoI	5343
Amorces de vérification de construction			
T7 promoterF	TAATACGACTCACTATAGGG		
T7 terminatorR	GCTAGTTATTGCTCAGCGG		

Les sites de restriction utilisés pour les clonages sont en rouge et soulignés dans chaque séquence

### II.3. RT-qPCR (Reverse Transcription quantitative Polymerase Chain Reaction)

La RT-qPCR (Reverse Transcription quantitative Polymerase Chain Reaction) a été réalisée comme suit : les amorces utilisées pour quantifier l'expression relative de gène ont été déterminées à l'aide du logiciel Primer3Plus, et listées dans le tableau 3 de l'article 4. L'absence de contamination par de l'ADN chromosomique a été vérifiée pour chaque échantillon avant la réaction de transcription inverse et un traitement par 3 U de DNase I RNase-free (Roche) a été réalisé le cas échéant. Les ADNc ont été synthétisés par le kit iScript Reverse Transcription Supermix de Bio-Rad selon les instructions du fabricant.

La même dilution d'ADNc a été testée trois fois en utilisant 5 µl d'ADNc (25 ng) qui ont été ajoutés à un mélange de PCR composé de 7,5 µl de SsoAdvanced™ Universal SYBR® Green Supermix et de 1,25 µl de chaque oligonucléotide à 0,8 µM. Les contrôles négatifs ont été réalisés en remplaçant 5 µl d'ADNc par 5 µl d'eau DEPC. Les PCR en temps réel (PCR quantitative ou qPCR) ont été réalisées dans un appareil CFX384 Real-Time thermocycler (BioRad). Le programme d'amplification est constitué de deux étapes : l'étape de dénaturation initiale à 95°C pendant 3 min et l'étape d'amplification de 38 cycles (chaque cycle est composé d'une étape de dénaturation de 15 s à 95°C et d'une étape d'hybridation et d'élongation de 30 s à 60°C). La mesure de fluorescence par un agent fluorescent non spécifique qui se fixe à l'ADN (SYBR® Green I) est réalisée à chaque étape d'hybridation. A la fin du programme, une étape supplémentaire de mesure de fluorescence entre 90°C et 60°C par incrément de 0,05°C a été réalisée pour établir une courbe de fusion afin de vérifier l'homogénéité des « amplicons ». L'efficacité de l'amplification a été déterminée en établissant une courbe standard avec une série des dilutions d'ADN. L'expression normalisée ( $\Delta\Delta C_T$ ) a été calculée par la méthode du seuil critique (Hellemans *et al.* 2007; Livak and Schmittgen 2001) dans laquelle la quantité relative (RQ) d'un échantillon pour un gène donné a été calculée par la formule :

$$RQ \text{ échantillon (gène)} = E_{\text{gène}}^{(Cq_{\text{(référence)}} - Cq_{\text{(échantillon)}})}$$

où E correspond à l'efficacité des amorces (cette efficacité est calculée avec la formule ((% d'efficacité x 0,01) + 1) ;  $Cq_{\text{(référence)}}$  correspond au  $Cq$  de l'échantillon de référence (ex : *L. casei* ATCC 334) pour le gène testé;  $Cq_{\text{(échantillon)}}$  correspond au  $Cq$  de l'échantillon testé (ex : mutant) pour le gène testé.

Cette quantité relative a permis de calculer l'expression normalisée en utilisant la formule :

$$\text{Expression normalisée échantillon (gène)} = \frac{RQ \text{ éch (gène)}}{(RQ \text{ éch (réf1)} \times RQ \text{ éch (réf2)} \times \dots \times RQ \text{ (éch réf n)})^{\frac{1}{n}}}$$

où RQ correspond à la quantité relative d'un échantillon (ex : mutant) ; réf correspond aux gènes de références.

Les gènes *recG* et *leuS*, dont l'expression est constante dans des conditions différentes, ont été choisis comme contrôles internes.

## II.4. PCR quantitative

De la même façon que la RT-qPCR, la PCR quantitative (qPCR) a été utilisée pour quantifier les mutants de la banque de *L. casei* grâce aux 70 différents tags. L'ADN des ensembles de mutants injectés et récupérés après séjour *in vivo* a été utilisé comme matrice pour déterminer la proportion de chaque mutant. Les amorces utilisées sont présentées dans le tableau S4 de l'article 3. L'amorce QPVITAGR est commune à tous les 70 couples d'amorce et la seconde amorce a été dessinée sur les 70 différents tags que peuvent posséder les mutants. Les couples d'amorces contrôles (CTRL3F et 3R, CTRL4F et 4R) ont été utilisés pour quantifier la totalité des tags. Les tags ont été quantifiés en utilisant la méthode du Cq. Les moyennes des valeurs de Cq obtenus avec les amplicons des amorces contrôles ont été utilisées pour déterminer le ratio de chaque tag sur la totalité des tags et cela pour les ensembles de mutants injectés et récupérés.

## II.5. Séquençage d'ADN

Les différentes constructions plasmidiques (partie IV) ont été vérifiées par séquençage chez GATC Biotech. La détermination du site d'insertion du transposon de la banque de mutants de *L. casei* a été réalisée, quant à elle, par séquençage en collaboration avec Magali Tichit et Christiane Bouchier de la plateforme génomique de l'Institut Pasteur, Paris. La technique de séquençage utilisée est développée dans l'article 2 des résultats « Rapid 96-well plates DNA extraction and sequencing procedures to identify genome-wide transposon insertion sites in a difficult to lyse bacterium: *Lactobacillus casei*. ».

## III. Techniques de transformation

### III.1. Transformation de *L. casei*

La souche *L. casei* a été transformée d'après la méthode de Keersmaecker *et al.* (De Keersmaecker *et al.* 2006). Une préculture a été utilisée pour ensemercer en cascade de  $10^{-2}$  à  $10^{-6}$  dans des tubes préchauffés contenant 5 ml de MRS supplémenté par 2% de glycine (m/v), les tubes ont été incubés 16 h à 37°C. Le tube pour lequel la  $DO_{600nm}$  est comprise entre 0,8 et 1 est utilisé, 2,5 ml de culture ont été prélevés et ajoutés à 50 ml de MRS + glycine 2% à 37°C et incubés. Après 3 h d'incubation de l'ampicilline a été ajoutée à 10 µg/ml. Après à nouveau 1 h d'incubation, les cellules ont été lavées deux fois dans du TSM (Tampon  $KH_2PO_4/K_2PO_4$  7 mM pH 7,4 ; 0,2 g.L<sup>-1</sup>  $MgCl_2$ , 6 H<sub>2</sub>O ; 171 g.L<sup>-1</sup> de saccharose) puis reprises dans 500 µl et laissées 30 min dans la glace.

La transformation des cellules a été effectuée par électroporation dans des cuves dont les électrodes sont espacées de 2 mm à l'aide d'un électroporateur (Genepulser, Biorad). Les cellules compétentes (100 µl) ont été mises en présence d'un échantillon de plasmide dessalé puis un pulse électrique de 1,7 kV, 200 Ohms et 25 µF a été appliqué. Les cellules ont été récupérées par au moins trois passages de 1 ml de MRS MC (MRS ; 2 mM CaCl<sub>2</sub> ; 20 mM MgCl<sub>2</sub>) et incubées dans un volume final de 5 ml, 3 h à 37°C. Des dilutions appropriées de la culture ont été étalées sur boîte contenant les antibiotiques de sélection afin d'obtenir des colonies de transformants.

### **III.2. Transformation d'*E. coli***

La souche *E. coli* TG1 a été transformée par électroporation (Gene pulser II Biorad, 2,5 kV, 200 Ohms, 25 µF) selon la méthode de Dower (Dower *et al.* 1988) et la souche *E. coli* BL21 (DE3) par transformation chimique (CaCl<sub>2</sub>) selon les instructions du fabricant (Invitrogen).

## **IV. Construction plasmidique**

En général, les constructions des plasmides ont suivi la procédure suivante : le fragment d'ADN à cloner a été obtenu par PCR (les amorces sont reportées dans le Tableau 7) puis purifié par le kit EZ-10 Spin Column PCR Product Purification de BioBasic Inc. Les fragments et le vecteur plasmidique obtenus ont été digérés par les enzymes de restriction identiques adéquates de chez New England, BioLabs<sup>®</sup>, Inc. Les ligations ont été réalisées à l'aide de l'ADN ligase T4 (New England, BioLabs<sup>®</sup>, Inc). Pour le premier clonage dans le vecteur pTZ19R (Tableau 8), la sélection des clones ayant intégré l'insert a été réalisée par le criblage LacZ sur milieu avec Xgal (criblage blanc-bleu). Pour les clonages dans les vecteurs pET28a+ (Tableau 8), la sélection des clones a été réalisée par PCR en utilisant les couples des amorces internes du plasmide vecteur (amorces T7).

Tableau 8 : Souches microbiennes, vecteurs et plasmides utilisés dans ce travail

Souche et plasmides	Principales caractéristiques	Source ou référence
<b>Souches de bases</b>		
<i>L. casei</i> ATCC 334	Type sauvage, Bacille Gram (+), CIP 107868	Collection Institut Pasteur, France
<i>E. coli</i> TG1	<i>SupE hsdΔ5thi Δ(lac-proAB)F' [tra D36 pro AB<sup>+</sup> lac I<sup>f</sup> lacZAM15]</i>	(Gibson 1984)
<i>E. coli</i> BL21 (DE3)	F <sup>-</sup> <i>ompT hsdSB</i> (rB mB <sup>-</sup> ) <i>gal dcm</i> (DE3)	Invitrogen
<b>Vecteurs de base</b>		
pTZ19R	Amp <sup>r</sup> , <i>AlacZ</i>	(Rokeach <i>et al.</i> 1988)
pET28a+	Kan <sup>r</sup> , vecteur pour la surexpression des protéines avec fusion his-tag, utilisant le promoteur du bactériophage T7	Novagen
<b>Plasmides</b>		
pET0219bis	pET28a+ avec le gène LSEI_0219 sous le promoteur T7 entre les sites de restrictions NcoI et XhoI permettant de surproduire le RR0219 tagué par histidines en C-terminal.	Ce travail
pET0220HAMP	pET28a+ avec une partie du gène LSEI_0220 (partie codant uniquement pour la partie cytosolique de la protéine) sous le promoteur T7 entre les sites de restrictions NcoI et XhoI permettant de surproduire la cHK0220 taguée par 6 histidines en C-terminal.	Ce travail
pTZ0219	pTZ19R contenant le gène LSEI_0219 ligué en extrémité franche grâce au site SmaI.	Ce travail
pET0219N	pET28a+ avec le gène LSEI_0219 sous le promoteur T7 entre les sites de restrictions NdeI et XhoI permettant de surproduire le RR0219 tagué par 6 histidines en N-terminal.	Ce travail

## V. Techniques de manipulation des protéines

### V.1. Obtention d'extrait brut acellulaire

#### V.1.1. Préparation des cellules

Le volume de cellules nécessaire a été récupéré par centrifugation à 3220 x g pendant 5 min à 4°C puis les cellules ont été lavées deux fois dans un tampon (NaCl 150 mM, PMSF 1mM, Tris-HCl 50 mM pH 7,5), centrifugées et reprises dans le volume final de 1 ml pour la casse au broyeur ou dans 5 ml pour la casse au disrupteur. Les extraits bruts acellulaires ont ensuite été obtenus par la casse des cellules aux billes de verre ou au disrupteur.

#### V.1.2. Casse des cellules au disrupteur

Les disruptions ont été réalisées avec le disrupteur de cellules Z Plus Series Cell Disrupter (Constantsystems). Les conditions de pression et le nombre de passages ont été adaptés pour

chaque microorganisme : un passage à  $10^8$  Pa pour *E. coli* et 2 passages à  $2,5 \cdot 10^8$  Pa pour *L. casei*.

### **V.1.3. Casse des cellules aux billes de verre**

La casse cellulaire a été réalisée dans un mini broyeur Fastprep 24 (MP Biomedicals). Les cellules ont été préparées comme auparavant et reprises dans 1 ml dans des tubes à FastPrep contenant des billes d'un diamètre moyen de 50  $\mu\text{m}$  (lysing matrix B, MPBio). Les protocoles de casse étaient les suivants : trois casses à puissance 5,5 pendant 15 s pour les extractions de protéines d'*E. coli* et de trois casses à puissance 5,5 pendant 20 s pour les extractions de protéines de *L. casei*. Les échantillons ont été mis pendant 1 min dans l'eau glacée entre les trois passages.

### **V.1.4. Extraction des protéines de surface liées de manière non covalente**

Les protéines de surface liées de manière non covalente sont extraites par la méthode au LiCl (Smit *et al.* 2001). Le volume de 20 ml d'une culture de 12 h a été récupéré par centrifugation à 3220 g pendant 5 min à  $10^\circ\text{C}$  puis les cellules ont été lavées deux fois dans de l'eau distillée à  $4^\circ\text{C}$ . Le culot cellulaire est ensuite repris dans 2 ml de LiCl 5M et incubé à  $37^\circ\text{C}$  sous agitation à 150 rpm pendant 1 h. Après centrifugation à 10000 g à  $10^\circ\text{C}$  pendant 20 min, les protéines du surnageant sont précipitées avec 4 ml d'acétone froid et incubées pendant 2 h à  $-20^\circ\text{C}$ . Les protéines ainsi précipitées sont récupérées par centrifugation (10000 g à  $4^\circ\text{C}$  pendant 10 min). Le précipité est resuspendu dans 100  $\mu\text{l}$  de tampon de charge (20  $\text{g.l}^{-1}$  de SDS, 8% (v/v) de glycérol, 0,1  $\text{g.l}^{-1}$  de Bleu de Bromophénol, 5 mM de DTT, 62,5 mM de tris-HCl) et chauffé 3 min à  $100^\circ\text{C}$  avant d'être analysé sur SDS-PAGE 12,5%.

## **V.2. Gels protéiques**

### **V.2.1.SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis)**

Les extraits protéiques ont été repris dans du tampon de charge (20  $\text{g.l}^{-1}$  de SDS, 8% (v/v) de glycérol, 0,1  $\text{g.l}^{-1}$  de Bleu de Bromophénol, 5 mM de DTT, 62,5 mM de tris-HCl), chauffés 3 min à  $100^\circ\text{C}$  afin de dénaturer les protéines et déposés sur un gel d'acrylamide à 12,5% (m/v) avec un marqueur de masse moléculaire (Bench Mark Ladder protein, Invitrogen). La migration a été effectuée à  $20 \text{ V.cm}^{-1}$  dans un tampon de migration à 1  $\text{g.l}^{-1}$  de SDS à l'aide du système Mini Protean 3 (Biorad). Puis le gel a été révélé dans une solution de bleu de Coomassie.

### V.2.2. BN-PAGE (Blue Native PAGE)

Afin d'étudier la phosphorylation du RR0219, des extraits RR0219 purifiés ont été ajoutés à du tampon de charge non dénaturant (8% (v/v) de glycérol, 0,1 g.l<sup>-1</sup> de Bleu de Bromophenol, 62,5 mM de Tris-HCl) et immédiatement chargés sur un gel d'acrylamide non dénaturant (sans SDS). Le taux d'acrylamide du gel et le pH du gel et de la solution de migration sont optimisés afin de permettre la migration la plus résolutive des protéines puisque la migration des protéines dans ce type de gel dépend de la masse moléculaire de l'oligomère protéique et du point isoélectrique. Dans le cadre de l'étude de la phosphorylation du RR0219, les extraits protéiques ont été résolus dans un gel à 12% (m/v) d'acrylamide à pH 8,8.

## V.3. Surexpression, purification et phosphorylation de protéines

### V.3.1. Surexpression de protéines

Dans le cadre de cette thèse, les protéines cHK0220 et RR0219 ont été surexprimées par le système d'expression du promoteur T7 du vecteur pET28a<sup>+</sup> (Biolabs). Ce vecteur possède un promoteur T7 inductible par l'IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) et deux séquences codant 6 résidus histidines (His-tag) en région N-terminale et C-terminale. Le fragment d'ADN correspondant au gène de la protéine à surexprimer a été amplifié par PCR à l'aide d'amorces qui créent des sites de restriction afin de permettre la ligation entre les sites NcoI et XhoI du plasmide pour avoir au final un His-tag en C-terminal et entre les sites NdeI et XhoI pour un His-tag en N-terminal (Tableau 7). Dans le cas de la surexpression de RR0219 avec un His-tag en C-terminal, le codon stop du gène a été remplacé par le site de restriction XhoI, ce qui a permis la création d'une fusion transcriptionnelle de 6 résidus histidine à la région C-terminale de la protéine (plasmide pET0219bis). Dans le cas de la surexpression de cHK0220 (partie cytosolique de la protéine HK0220), une partie du gène a été clonée, cette partie de gène ne codant que pour la partie cytosolique de la protéine, le codon stop a aussi été remplacé par le site de restriction XhoI, ce qui a permis la création d'une fusion transcriptionnelle de 6 résidus histidine à la région C-terminale de la protéine (plasmide pET0220HAMP). Dans le cas de la surexpression du RR0219N avec un His-tag en N-terminal, l'intégralité du gène a tout d'abord été sous-clonée dans le vecteur pTZ donnant ainsi le plasmide PTZ0219, puis l'intégralité du gène a été clonée dans de pET28a<sup>+</sup> permettant une surexpression de la protéine originelle chez *E. coli* (plasmide pET0219N). Dans tous les cas, la fusion en 5' place le gène d'intérêt sous le contrôle du promoteur T7. Les plasmides ainsi obtenus ont été transformés chez *E. coli* BL21 (DE3) qui possède le gène de la T7 ARN polymérase sous contrôle du promoteur inductible *lacUV*. L'ajout d'IPTG permet l'expression de la T7 ARN polymérase qui transcrit le gène cloné.

### V.3.2. Purification de protéines

Les cellules provenant d'un litre de culture de la souche recombinante *E. coli* pET0219bis, induites par 0,5 mM d'IPTG (Euromedex) ajouté à une DO<sub>600</sub> de 0,4 et incubé pendant une nuit, ont été obtenues par centrifugation, lavées puis reprises dans 10 mL de tampon (NaCl 150 mM, Tris-HCl, 50 mM, pH7,5). Les extraits bruts ont ensuite été obtenus par casse au disrupteur CellD Constant ou par broyage au FastPrep, puis chargés sur une colonne de 1 mL de nickel-nitrilotriacetate (Ni-NTA) (Macherey-Nagel). Grâce à l'affinité de cette matrice vis-à-vis des résidus histidines (His-tag), RR0219 est retenue à la colonne. Une série de 3 lavages permet de « décrocher » les protéines retenues de façon non spécifique. L'élution de RR0219, se fait par un compétiteur des résidus histidines vis-à-vis de la colonne Ni-NTA, l'imidazole ajouté en gradient de 80 à 250 mM. La purification des protéines RR0219N et cHK0220 a été réalisée de la même façon à partir des souches recombinantes *E. coli* pET0219N et *E. coli* pET0220HAMP, respectivement. La qualité de la purification des protéines a été vérifiée en électrophorèse sur gel SDS-PAGE à 12,5%. La concentration des protéines a été déterminée par l'utilisation du kit de dosage de protéines Bio-Rad avec l'albumine de sérum bovin comme standard.

### V.3.3. Clivage du His-tag

Pour la cadre des EMSAs réalisés lors de cette thèse, la protéine RR0219N (His-tag en N-terminal) a été produite pour remplacer la protéine RR0219 (His-tag en C-terminal) ; en effet, la protéine RR0219 ne se fixant pas sur la sonde, la protéine RR0219N a été produite afin d'être sûr qu'il ne s'agissait pas d'un encombrement stérique de l'étiquette histidine au niveau du domaine de fixation. La protéine RR0219N est constituée de N- en C-terminal d'une étiquette histidine, d'un site de clivage de la thrombine (Leu-Val-Pro-Arg-Gly-Ser) et de la protéine d'intérêt. L'étiquette histidine a été clivée grâce à la présence du site de clivage de la thrombine, le kit « thrombin CleanCleave » (Sigma) a été utilisé pour cette manipulation en suivant les instructions du fabricant. Le clivage des résidus d'acides aminés reconnu par la thrombine a ensuite été vérifié sur un gel SDS-PAGE à 12,5%.

### V.3.4. Autophosphorylation et transfert de phosphate

Afin de vérifier la fonctionnalité du système à deux composants étudié lors de cette thèse, la phosphorylation de l'HK puis le transfert du phosphate de cette dernière vers son RR ont été réalisés ainsi que l'autophosphorylation du RR à partir d'une petite molécule donneuse de phosphate.

Pour la réaction d'autophosphorylation, 6 µg de cHK0220 ont été incubés dans un volume final de 20 µl de tampon de phosphorylation (50 mM Tris-HCl pH 7.5, 2,5 mM MnCl<sub>2</sub>, 5 mM

MgCl<sub>2</sub>, 50 mM KCl) auquel a été ajouté 20 µCi of [ $\gamma$ -<sup>32</sup>P]- ATP (3000 Ci/mmol, Perkin Elmer) pendant 1, 5, 15, 30 et 60 min à 37°C. La réaction a été initialisée par l'ajout d'ATP radioactif et stoppée par l'ajout de tampon de dépôt contenant du SDS. Pour la réaction de transfert de phosphate, 10 µg de RR0219 ont été ajoutés après 60 min d'autophosphorylation de la cHK0220 et des aliquotes ont été récupérées après 0, 15 et 60 min d'incubation à 37°C. La réaction a été stoppée par l'ajout de tampon de dépôt contenant du SDS. Les différents échantillons ont été analysés sur gel SDS-PAGE 12,5% puis révélés par autoradiographie.

L'autophosphorylation de la protéine RR0219 ou RR0219N par une petite molécule donneuse de phosphate a été vérifiée de deux façons différentes. Pour la première méthode, du [<sup>32</sup>P] acetyl-phosphate a été préparé en incubant pendant 30 min à 37°C de l'acetate kinase (0,4 units, Sigma) dans 55 µl de tampon contenant 25 mM de Tris-HCl pH 7,5, 10 mM de MgCl<sub>2</sub> et de 60 mM de potassium acetate ainsi que du [ $\gamma$ -<sup>32</sup>P]- ATP (3000 Ci/mmol, Perkin Elmer). Le [<sup>32</sup>P]-acetyl phosphate ainsi préparé a ensuite été incubé à 37°C, à différents temps, avec la protéine RR0219 dans 40 µl dans le tampon de phosphorylation. La réaction a été arrêtée par ajout de tampon de dépôt contenant du SDS. Les échantillons ont été analysés sur gel SDS-PAGE 12,5% puis révélés par autoradiographie. La seconde méthode a consisté à incuber à 37°C la protéine RR0219 (5 µg) ou RR0219N (5 µg) avec 60 mM de l'acetyl-phosphate dans le tampon de phosphorylation à des temps différents et de suivre la phosphorylation sur un gel natif à 12,5% de polyacrylamide (BN-PAGE). La protéine phosphorylée a migré plus loin que la protéine non phosphorylée car chargée plus négativement. Cette seconde méthode a principalement été utilisée à chaque fin de purification pour vérifier la fonctionnalité de phosphorylation de la protéine avant de réaliser des EMSAs (Electrophoretic Mobility Shift Assay).

## **VI. EMSA (Electrophoretic Mobility Shift Assay)**

La technique d'EMSA, aussi nommée retard sur gel, a pour objectif de mettre en évidence des interactions spécifiques entre des protéines et un fragment d'ADN cible correspondant le plus souvent à une région promotrice. Si une protéine se fixe de manière spécifique à un fragment d'ADN, il se forme un complexe de masse moléculaire plus élevée que le fragment d'ADN seul. Lors de la migration sur gel de polyacrylamide, le complexe protéine-ADN migrera moins que la sonde ADN seule d'où l'appellation de retard sur gel.

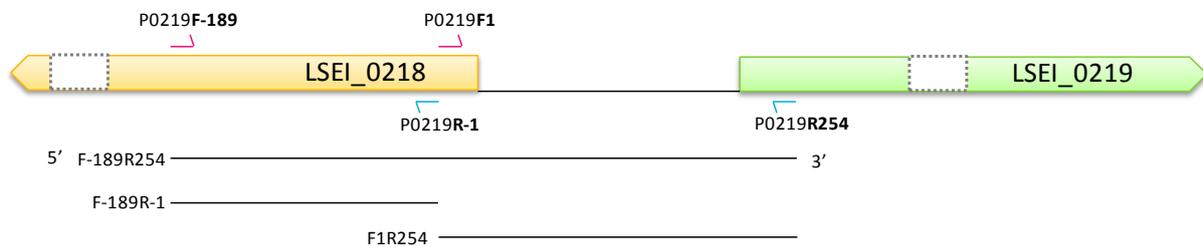
Les sondes ADN sont marquées en extrémité 3' avec la digoxigénine. Dans un premier temps, la sonde est amplifiée par PCR d'après le protocole standard (voir la partie II.2) à l'aide des amorces présentées dans le Tableau 9, puis purifiée avec une mini colonne de purification, kit EZ-10 Spin Column PCR Product Purification kit, BioBasic Inc ou MinElute Reaction Cleanup Kit, Qiagen, suivant la taille de la sonde. Le marquage de la sonde a ensuite été réalisé avec le kit DIG Gel Shift Kit, 2<sup>nd</sup> Generation. La ddUTP-11-digoxigénine est liée à

l'extrémité 3' de la sonde grâce à une transférase terminale. Les différentes sondes utilisées sont reportées dans la Figure 32.

La réaction de fixation a eu lieu à 30°C, pendant 20 min, dans le volume de 15 µL contenant le tampon de fixation (Tris-HCl pH 7,5, 15 mM ; glycérol, 5% (v/v) ; EDTA, 0,2 mM ; KCl, 25 mM ; NaCl, 25 mM ; MgCl<sub>2</sub>, 2,5 mM ; DTT, 2 mM ; BSA, 2,5 µg.ml<sup>-1</sup> ; ADN de sperme de saumon, 2,5 µg.ml<sup>-1</sup>), la sonde ADN marquée (~ 10<sup>-7</sup> M), extrait protéique (concentration variable précisée dans l'article 4) chauffé 10 min à 65°C. Le mélange est chargé sur un gel de polyacrylamide 6%, la migration a été effectuée dans un tampon contenant 20 mM tris-HCl, 400 mM glycine, 1 mM EDTA, pH 7,8 à 80 V et en chambre froide pendant 1h30. Enfin le gel a été transféré sur une membrane de nylon chargée positivement (Roche) puis les sondes ont été révélées par chimioluminescence selon le protocole du kit DIG Gel Shift Kit, 2<sup>nd</sup> Generation.

**Tableau 9 : Amorces utilisées pour synthétiser les différentes sondes.**

Nom sonde		Nom amorce	Séquence 5'→ 3'	taille (pb)	
F-189R254	F-189R-1	P0219F-189	GCTAAGGCTTTTGTC AATATGG	189	443
		P0219R-1	CATCGCTGGCATT AACATTGC		
	F1R254	P0219F1	TGCATGGTTGAGGGTAAACT	254	
		P0219R254	CTTAGTAATTCGACGATTTC		



**Figure 32 : Localisation des différentes sondes utilisées pour réaliser les EMSAs.**

Les gènes ne sont pas représentés en intégralité.

## VII. Autolyse induite au Triton X-100

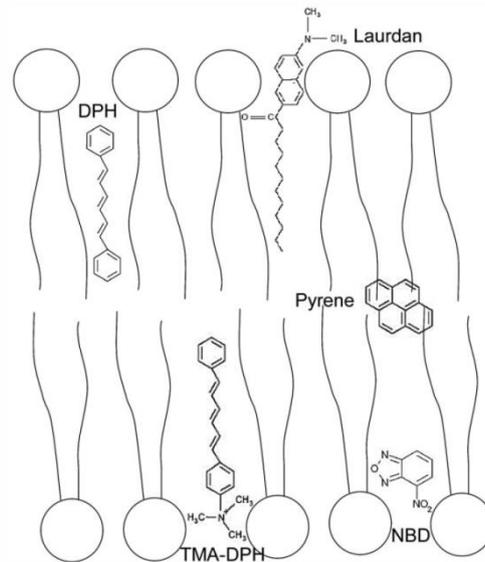
Des cultures en phase exponentielle (DO<sub>600nm</sub> de 0,8) ont été centrifugées et lavées dans du tampon phosphate de potassium 50 mM pH 7,0 puis resuspendues à une DO<sub>600nm</sub> de 1 dans le même tampon phosphate (contrôle) ou dans un tampon phosphate contenant du Triton X-100 à 0,05% (Cornett and Shockman 1978). Ces suspensions ont été incubées à 30°C pendant 24 h. L'autolyse a été déterminée par la diminution de la DO<sub>600nm</sub>. Le pourcentage d'autolyse est ensuite calculé en divisant le DO<sub>600nm</sub> mesurée par rapport à la DO<sub>600nm</sub> initiale.

## VIII. Mesure de la variation de la fluidité membranaire

La définition de la fluidité membranaire se réfère aux propriétés dynamiques des lipides dans la bicouche membranaire et, par extension, à leurs mouvements moléculaires et désordonnés. Un changement dans les caractéristiques structurales et dynamiques de la membrane va affecter les fonctions des protéines membranaires par la modification de leur environnement lipidique. Or un fonctionnement cellulaire correct nécessite une fluidité membranaire optimale. Ceci explique la nécessité d'une adaptation homéovisqueuse : des perturbations externes (température, pH, agent chimique...) qui affecteraient la fluidité membranaire sont compensées de sorte à maintenir une fluidité constante (Sinensky 1974).

Plusieurs sondes fluorescentes peuvent être utilisées suivant leur localisation lors de l'incorporation dans la membrane (Figure 33). Pour évaluer la fluidité membranaire de *L. casei*, nous avons choisi la sonde hydrophobe 1,6-diphenyl-1,3,5-hexatriene (DPH) (Trevors 2003). Le DPH s'aligne avec les chaînes acylées des phospholipides et est distribué au centre de la bicouche phospholipidique. L'anisotropie de fluorescence de la sonde DPH fournit des données sur le comportement moyen des lipides membranaires suivant les mouvements de rotation de la sonde dans la bicouche membranaire.

Les travaux précédemment menés au laboratoire ont permis d'établir un protocole de mesure de fluidité membranaire sur des cellules bactériennes entières, ce qui présente l'avantage de mesurer les variations instantanées de la fluidité membranaire de cellules stressées au cours de la mesure. Les cellules de *L. casei* ont été cultivées en conditions classiques jusqu'à une absorbance à 600 nm de 0,6, centrifugées, lavées deux fois et reprises dans un volume de tampon phosphate 25 mM (pH 6,0) supplémenté avec 10 mM de glucose. Trois ml de suspension cellulaire ont été placés dans une cuvette en quartz agitée et thermostatée à 37°C grâce à un système Peltier (Wavelength Electronics). La concentration de la sonde DPH (Sigma) était de 2 µM.



**Figure 33 : Sondes fluorescentes permettant d'évaluer la fluidité membranaire.**

La bicouche est schématisée par deux couches phospholipidiques opposées, arrangées de manière à ce que les têtes des phospholipides (régions hydrophiles) soient en contact avec l'eau et que les chaînes d'acides gras (régions hydrophobes) se fassent face. Différentes sondes peuvent être utilisées pour déterminer les propriétés de la membrane selon leurs caractéristiques. Les sondes polaires comme la NBD (nitrobenzoxadiazole) ou les sondes chargées comme le TMA-DPH sont souvent utilisées pour examiner l'interface membrane-eau. Le laurdan est utilisé pour étudier la polarité membranaire. Le DPH ou le pyrène, très hydrophobes, sont couramment utilisés pour explorer la région des chaînes d'hydrocarbone de la membrane. D'après (Cao-Hoang *et al.* 2008).

L'anisotropie de fluorescence de la sonde est inversement proportionnelle à la fluidité membranaire. Les intensités de lumière polarisée parallèles et perpendiculaires au plan de polarisation du faisceau d'excitation permettant la détermination de l'anisotropie de fluorescence ont été mesurées avec un spectrofluorimètre sous format T (Fluorolog-3, Jobin Yvon). L'intensité de fluorescence de la sonde DPH a été mesurée à 431 nm pour une excitation à 360 nm. L'anisotropie de fluorescence ( $r$ ) est calculée selon l'équation suivante (Lakowicz and Thompson 1983) :

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}}$$

$I_{VV}$  et  $I_{VH}$  représentent respectivement les intensités mesurées avec les polariseurs parallèles et perpendiculaires au faisceau d'excitation vertical de lumière polarisée.  $G$  est le facteur utilisé pour corriger la polarisation de l'appareil, il est donné par le ratio des composantes verticale sur horizontale de l'émission polarisée quand la lumière d'excitation est polarisée dans la direction horizontale.

## IX. Microscopie électronique à transmission (TEM)

(Travail réalisé en collaboration avec Jeannine Lherminier de la plateforme de microscopie de l'INRA de Dijon)

La microscopie électronique à transmission (TEM) a été utilisée au cours de cette thèse afin de déterminer l'épaisseur de la paroi de *L. casei* ATCC 334 et d'un mutant. Cette méthode consiste en la fixation des bactéries en suspension suivie d'une inclusion en résine époxy (EmBed 812/DER 736) permettant de réaliser des coupes ultrafines pour des études morphologiques ultra structurales. Les coupes ont été marquées par la méthode de PATAg (Periodic Acid-Thiocarbaside-Protéinate d'Argent) pour mettre en évidence des polysaccharides (Thiéry 1967). Les observations ont été faites avec un microscope électronique de transmission Hitachi 7500 à 80 kV équipé d'une caméra CCD à haute résolution AMT Advantage. Les images transmises ont été enregistrées et traitées par le logiciel AMT.

## X. Tests d'adhésion

### X.1. Adhésion microbienne aux solvants (MATS, Microbial Adhesion To Solvent)

Cette méthode, basée sur la comparaison de l'affinité des micro-organismes pour les solvants polaires et apolaires, de tensions de surface connues permet d'évaluer les caractères électron-donneur ou électron accepteur des micro-organismes et révéler ainsi les propriétés hydrophiles ou hydrophobes (Bellon-Fontaine *et al.* 1996). L'adhérence microbienne est liée aux caractéristiques de surface de la cellule, notamment hydrophobicité et la charge.

Afin de déterminer si une différence d'hydrophobicité et de charge existe entre les mutants sélectionnés et *L. casei*, des tests d'adhésion à différents solvants ont été réalisés. Pour cela, les bactéries ont été récupérées en phase stationnaire de croissance, lavées en tampon PBS pH 7,4 et la DO<sub>600 nm</sub> de la suspension a été ajustée à 0,4 (A<sub>0</sub>). Un volume de 3 ml de suspension bactérienne a ensuite été ajouté à 1 ml de solvant (chloroforme, hexadécane ou acétate d'éthyle) et mélangé au vortex pendant 3 minutes. Après 15 minutes de repos afin de permettre la séparation des phases aqueuse et organique, la DO<sub>600 nm</sub> de la phase aqueuse est mesurée (A<sub>1</sub>). Le pourcentage d'adhésion est ensuite calculé selon la formule :

$$\% \text{ adhésion} = 100 \times \left( \frac{A_0 - A_1}{A_0} \right)$$

Le chloroforme est un solvant acide et accepteur d'électron, l'hexadécane est un solvant hydrophobe et enfin l'acétate d'éthyle est un solvant basique et donneur d'électron (Bellon-Fontaine *et al.* 1996).

## X.2. Adhésion à la mucine

L'adhésion à la mucine a été testée afin de savoir si une différence existait entre la souche de *L. casei* et les mutants sélectionnés. Des plaques 6 puits non traitées en polystyrène (ThermoScientific, Nunc) ont été recouvertes de mucine d'estomac de porc type II à 0,5 g.l<sup>-1</sup> (Sigma) une nuit à 4°C puis lavées 2 fois avec du PBS pH 7,4. La DO<sub>600nm</sub> de culture de 16 h en milieu Lactobacilli AOAC (Difco) a été ajustée à 1. 2 ml de chaque culture ont été mis en contact avec la mucine dans les plaques 6 puits pendant 3h30 à 37°C. Les plaques ont ensuite été lavées 2 fois avec du PBS pH 7,4 afin d'éliminer les bactéries n'ayant pas adhéré, puis le fond des puits a été gratté avec des spatules Cell Lifter (Dutscher) et les bactéries ayant adhéré ont été dénombrées sur boîte MRS pour *L. casei* et sur MRS Ery 5 µg.ml<sup>-1</sup> pour les mutants après avoir réalisé une série de dilution. L'inoculum a aussi été dénombré sur boîte au bout du temps d'incubation en tube. Le pourcentage d'adhésion est ensuite calculé selon la formule :

$$\% \text{ adhésion} = 100 \times \left( \frac{N_{\text{adhérées}}}{N_{\text{inoculum}}} \right)$$

où  $N_{\text{adhérées}}$  correspond à la concentration dénombrée des bactéries ayant adhéré à la mucine et  $N_{\text{inoculum}}$  à la concentration dénombrée après incubation des bactéries en tube.

## X.3. Adhésion à une surface abiotique

L'adhésion à une surface abiotique a été testée afin de savoir si une différence existait entre la souche de *L. casei* et les mutants sélectionnés. La DO<sub>600nm</sub> de culture de 16 h en milieu Lactobacilli AOAC (Difco) a été ajustée à 0,1 et 2 ml de chaque culture ont été mis en contact avec des plaques 6 puits non traitées en polystyrène (ThermoScientific, Nunc) pendant 48 h à 37°C. Les plaques ont ensuite été lavées 2 fois avec du PBS pH 7,4 afin d'éliminer les bactéries n'ayant pas adhéré, puis le fond des puits a été gratté avec des spatules Cell Lifter (Dutscher) et les bactéries ayant adhéré ont été dénombrées sur boîte MRS pour *L. casei* et sur MRS Ery 5 µg.ml<sup>-1</sup> pour les mutants après avoir réalisé une série de dilutions. L'inoculum a aussi été dénombré sur boîte au bout du temps d'incubation. Le pourcentage d'adhésion est ensuite calculé selon la même formule que précédemment:

$$\% \text{ adhésion} = 100 \times \left( \frac{N_{\text{adhérées}}}{N_{\text{inoculum}}} \right)$$

où  $N_{\text{adhérées}}$  correspond à la concentration dénombrée après incubation des bactéries ayant adhéré à la surface abiotique et  $N_{\text{inoculum}}$  à la concentration dénombrée après incubation des bactéries en tube.

## XI. Compétition souche parentale et mutants

### XI.1. Compétition *in vivo* dans les anses iléales de lapin

Lors de cette thèse, la banque de mutant de *L. casei* réorganisée en une banque ne contenant que les mutants avec insertion du transposon dans une séquence génique a été testée dans des anses iléales ligaturées de lapin (Marteyn *et al.* 2010; Sansonetti *et al.* 1983; West *et al.* 2005). Une suspension bactérienne composée de 40 à 70 mutants différents et contenant  $5.10^7$  UFC a été injectée dans chaque anse. Après 16 h de compétition *in vivo*, la totalité de l'anse iléale est récupérée, homogénéisée, diluée puis étalée sur boîte MRS Ery  $5 \mu\text{g.ml}^{-1}$  afin d'obtenir, après 48 h d'incubation à  $37^\circ\text{C}$ , des colonies isolées. L'ADN est ensuite extrait des colonies. Des PCR quantitatives ont été réalisées sur les extraits d'ADN afin de déterminer la proportion de chaque mutant avant et après séjour *in vivo*. Le ratio a par la suite été calculé afin de déterminer la diminution de chaque mutant.

### XI.2. Compétition *in vitro* dans différents milieux

Afin de déterminer la sensibilité des mutants sélectionnés par rapport à *L. casei*, des compétitions dans différentes conditions stressantes ont été réalisées, avec un pourcentage initial de mutant de 2% et après une incubation de 16 h à  $37^\circ\text{C}$ . Les milieux testés sont :

- un milieu appauvri et acide (MRS dilué au 1/20 dans du PBS pH 3,5),
- un milieu appauvri et supplémenté en bile (MRS dilué au 1/20 dans du PBS supplémenté en bile à  $3 \text{ g.l}^{-1}$ ),
- un milieu appauvri et supplémenté en lysozyme (MRS dilué au 1/20 dans du PBS et supplémenté en lysozyme à  $10 \text{ g.l}^{-1}$ ),
- un milieu MRS supplémenté en pénicilline à  $0,25 \mu\text{g.ml}^{-1}$ ,

Les pourcentages du mutant par rapport à *L. casei* ont été déterminés par dénombrement sur boîte et par la formule suivante :

$$\% \text{ mutant} = 100 \times \left( \frac{N_{\text{mutant}}}{N_{L. \text{ casei}}} \right)$$

où  $N_{\text{mutant}}$  correspond à la concentration dénombrée du mutant et  $N_{L. \text{ casei}}$  à celle de *L. casei*.

Le ratio du % du mutant dans la condition testée par rapport à la condition initiale a ensuite été calculé pour déterminer la sensibilité des mutants sélectionnés.

## XII. Sensibilité aux antibiotiques

La sensibilité aux antibiotiques agissant sur la paroi tels que la pénicilline, la bacitracine, la nisine et la vancomycine a été testée sur *L. casei* et sur les mutants de l'opéron. Des plaques 96 puits contenant 10 µl d'antibiotique à différentes concentrations ont été inoculées avec 90 µl d'une suspension bactérienne à une  $DO_{600\text{nm}}$  de 0,01 dans du MRS. Après 24 h à 37°C, la croissance bactérienne visible a été observée et la plus faible concentration d'antibiotique qui a empêché la croissance a été déterminée comme étant la CMI (Concentration Minimale Inhibitrice).

## XIII. Analyse et traitement des données bioinformatiques

Les oligonucléotides ont été conçus à l'aide des logiciels Primer3 (<http://frodo.wi.mit.edu/>) et oligo Analyzer (<http://oligo-analyzer.software.informer.com/>). L'alignement multiple des séquences d'ADN ou de protéine a été réalisé à l'aide des logiciels Blast, Basic Local Alignment Search Tool (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) et ClustalW (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). La séquence complète du génome de *L. casei* ATCC 334 a été consultée dans le Genbank par numéro d'accèsion NC\_008526. Les logiciels biocyc (<http://biocyc.org/>) et KEGG (<http://www.genome.jp/kegg/>) ont été utilisés pour la prédiction fonctionnelle des protéines.

# Résultats

Cette partie « Résultats » rassemble les résultats des travaux que j'ai réalisés et ceux auxquels j'ai participé lors de ma thèse. Cette partie est présentée sous forme d'articles (acceptés ou en préparation) précédés d'un résumé en français et certains points, moins aboutis, sont présentés en français sous une forme condensée. Le plan, correspondant à la suite chronologique du travail de recherche, est le suivant :

- I. Mutagenèse par transposition de *L. casei* et identification des sites d'insertion du transposon
  - I.1. Mutagenèse par transposition de *L. casei*

Article 1. Development of an efficient in vivo system ( $P_{\text{junc}}$ -TpaseIS<sub>1223</sub>) for random transposon mutagenesis of *Lactobacillus casei*.  
Licandro-Seraut H, Brinster S, van de Guchte M, **Scornec H**, Maguin E, Sansonetti PJ, Cavin J-F and Serror P. 2012. Appl. Environ. Microbiol. 78: 5417-5423.
  - I.2. Identification des sites d'insertion du transposon des mutants de *L. casei*

Article 2. Rapid 96-well plates DNA extraction and sequencing procedures to identify genome-wide transposon insertion sites in a difficult to lyse bacterium: *Lactobacillus casei*.  
**Scornec H**, Tichit M, Bouchier C, Pédrón T, Cavin J-F, Sansonetti PJ and Licandro-Seraut H. 2014. J. Microbiol. Methods. 106: 78-82
- II. Identification des gènes impliqués lors de l'établissement de *L. casei* dans l'intestin.

Article 3. Functional genomics of *Lactobacillus casei* establishment in the gut.  
Licandro-Seraut H, **Scornec H**, Pédrón T, Cavin J-F and Sansonetti PJ. 2014. Proc. Natl. Acad. Sci. USA. 111 (30): E3101-9
- III. Caractérisation fonctionnelle de l'opéron LSEI\_0219-0221

Article 4. The LSEI\_0219-0221 operon: a central interface of the surface architecture of *Lactobacillus casei* which conditions its resistance to the gut.  
**Scornec H**, Pédrón T, Wheeler R, Boneca IG, Sansonetti PJ, Cavin J-F and Licandro-Seraut H. (En préparation)

## I. Mutagenèse par transposition de *L. casei* et identification des sites d'insertion du transposon

### I.1. Mutagenèse par transposition de *L. casei*

Ce premier article est le fruit d'une collaboration avec l'unité Micalis de l'INRA de Jouy-en-Josas. La mutagenèse de *Lactobacillus delbrueckii* subsp. *bulgaricus* a été réalisée par leur équipe alors que celle de *L. casei* a été réalisée par la nôtre. Ce travail a été réalisé en majeure partie avant le début de ma thèse, j'ai pu y apporter ma contribution lors du stage de master Recherche effectué au sein de l'équipe.

#### Principaux résultats de l'article 1.

*L. casei* est une bactérie Gram positif difficile à transformer ; en effet, cette bactérie a un taux de transformation qui est assez faible ( $10^2$  transformants par  $\mu\text{g}$ , taux variable en fonction des vecteurs) malgré les différentes techniques de transformation testées. Par conséquent peu d'études de génétique inverse ont été réalisées. L'article, ci-après, présente un système de mutagenèse par transposition efficace pour *L. casei*. Ce système consiste en l'utilisation de deux vecteurs fonctionnant en trans. Le premier vecteur, le pVII29, est thermosensible et comporte un gène codant pour la transposase IS1223 de *Lactobacillus johnsonii*, et le second vecteur, le pVII10, est suicide et possède les deux séquences inversées répétées ( $P_{\text{junc}}$ ) correspondant au substrat de la transposase.

Ces deux vecteurs ont été consécutivement transformés chez *L. casei*. Les transformants possédant le pVII29, nommés LCpVII29 ont été criblés quant à leur capacité à résister au chloramphénicol (le pVII29 possède le gène de résistance au chloramphénicol). Ces transformants ont à leur tour été transformés avec le vecteur pVII10 et criblés quant à leur capacité à résister à l'érythromycine. Le taux de transformation du LCpVII29 atteint alors entre 2,5 et  $7,5 \cdot 10^3$  par  $\mu\text{g}$  de pVII10. Le pVII29 va permettre, grâce à la reconnaissance par la transposase des séquences cibles, l'intégration chromosomique du pVII10. Les transformants se développant sur érythromycine sont considérés comme étant des intégrants génomiques (chromosome ou plasmide). Plusieurs générations de culture des intégrants sont ensuite réalisées pour éliminer le vecteur thermosensible pVII29, limitant ainsi l'instabilité des mutants.

Une analyse de l'insertion du pVII10 sur le génome de *L. casei* des mutants a ensuite été réalisée afin de déterminer le caractère aléatoire et singulier de la transposition ; pour cela, l'hybridation « Southern » entre l'ADN de mutant digéré par HindIII et une sonde spécifique du pVII10 a été testée. Il s'avère que l'intégration s'est faite dans un locus unique excepté pour 2 sur 12 mutants testés. Le séquençage de la région d'insertion a ensuite été réalisé sur 20 mutants et il s'avère que 50% des insertions se situent dans une région intergénique, 10%

se situent dans le plasmide et seulement 20% se situent dans une épingle à cheveu prédite. De plus, aucune insertion préférentielle dans les régions riches en A/T n'a été observée. Exceptées les deux insertions identiques dans le plasmide, les autres insertions ne sont présentes qu'une seule fois sur les 20 mutants testés, ceci laisse supposer un caractère suffisamment aléatoire de la mutagenèse par transposition pour créer une banque de mutants saturée.

Le système de mutagenèse présenté dans cet article permet d'obtenir un haut taux d'intégrants stables (10000 mutants par transformation de *L. casei*) malgré le faible taux de transformants des lactobacilles. Ce système peut être utilisé sur d'autres bactéries Gram positif et en particulier sur les bactéries lactiques.

## Development of an Efficient *In Vivo* System ( $P_{\text{junc}}$ -T $\text{paseIS}_{1223}$ ) for Random Transposon Mutagenesis of *Lactobacillus casei*

Hélène Licandro-Seraut,<sup>a,b</sup> Sophie Brinster,<sup>c,d,\*</sup> Maarten van de Guchte,<sup>c,d</sup> Hélène Scornec,<sup>a</sup> Emmanuelle Maguin,<sup>c,d</sup> Philippe Sansonetti,<sup>b</sup> Jean-François Cavin,<sup>a</sup> and Pascale Serror<sup>c,d</sup>

UMR PAM, AgroSup Dijon et Université de Bourgogne, Dijon, France<sup>a</sup>; Unité de Pathogénie Microbienne Moléculaire, Institut Pasteur, Paris, France<sup>b</sup>; INRA, UMR1319 Micalis, Jouy-en-Josas, France<sup>c</sup>; and AgroParisTech, UMR Micalis, Jouy-en-Josas, France<sup>d</sup>

The random transposon mutagenesis system  $P_{\text{junc}}$ -T $\text{paseIS}_{1223}$  is composed of plasmids pVI129, expressing *IS1223* transposase, and pVI110, a suicide transposon plasmid carrying the  $P_{\text{junc}}$  sequence, the substrate of the *IS1223* transposase. This system is particularly efficient in *Lactobacillus casei*, as more than 10,000 stable, random mutants were routinely obtained via electroporation.

Lactic acid bacteria (LAB) are widely used in food fermentations, as well as for their probiotic properties. *Lactobacilli delbrueckii* subsp. *bulgaricus* and *Lactobacillus casei* have been shown to provide beneficial effects to the immune system (29, 35). However, due to the lack of reliable tools such as a random mutagenesis system to perform global reverse genetics, the overall mechanisms underlying their probiotic effects are poorly understood.

Neither the Gram-positive transposition systems based on transposon delivery by a suicide or a thermosensitive vector (19, 27, 34, 36) nor *in vitro* transposon mutagenesis using Tn5-based transposons (17) is adapted to all species of LAB, due to low transformation efficiencies or unwanted stability of the transposon delivery vector (23).

IS3 sequences are surrounded by imperfect inverted repeats (IR). They carry two consecutive and partially overlapping open reading frames, *orfA* and *orfB*, which encode a transposase. IS3 sequences undergo a “cut-and-paste” transposition mechanism that occurs by generating a covalently closed circular transposition intermediate, which promotes transposase induction resulting from the generation of a strong promoter named  $P_{\text{junc}}$ . The  $P_{\text{junc}}$  promoter corresponds to abutted IRR (inverted repeat right) and IRL (inverted repeat left) sequences as a result of insertion sequence circularization and constitutes an efficient transposition substrate (13, 14). Here, we report the construction of a novel *in trans* transposition procedure, named the  $P_{\text{junc}}$ -T $\text{paseIS}_{1223}$  system and dedicated to *in vivo* random mutagenesis in LAB, and its application for random mutagenesis in *L. casei*. It is based on *IS1223*, a member of the IS3 family from *Lactobacillus johnsonii* (39) that transposes efficiently in *Lactobacillus delbrueckii* subsp. *bulgaricus* (31, 39). This system is composed of two plasmids: pVI129, carrying the *IS1223* transposase gene, and pVI110, a suicide transposon plasmid carrying the  $P_{\text{junc}}$  sequence, the substrate of the *IS1223* transposase.

**Construction of the  $P_{\text{junc}}$ -T $\text{paseIS}_{1223}$  system and validation in *Escherichia coli*.** Plasmid pVI116 was constructed as described in Fig. 1 and its legend to provide the transposase of *IS1223* expressed under the control of the *L. delbrueckii* subsp. *bulgaricus*  $P_{\text{hlbA}}$  promoter (9). Plasmid pVI115 was constructed from pVI162 (see Table 1 for details of construction) to provide the transposition substrate corresponding to an abutted IRR-IRL junction of *IS1223* separated by 3 base pairs, named  $P_{\text{junc}}$  (Fig. 1A). It replicates only in the TG1 RepA strain of *Escherichia coli* (18). Plasmid

pVI116 ( $P_{\text{hlbA}}$ -T $\text{paseIS}_{1223}$ ) and the control plasmids, pGB2 and pVI113 (T $\text{paseIS}_{1223}$ ), were electroporated into *E. coli* TG1 as previously described (12). The resulting strains were electroporated with identical amounts (100 ng) of pVI115, as a nonreplicative source of the  $P_{\text{junc}}$ , or pVI119 (Table 1), as a  $P_{\text{junc}}$ -less nonreplicative control, and pGEMT, as a replicative plasmid. Cells were directly plated on LB agar plates supplemented with chloramphenicol (10  $\mu\text{g}/\text{ml}$ ) or ampicillin (50  $\mu\text{g}/\text{ml}$ ). Plates were incubated for 20 h at 37°C, and colonies were counted to score for integration or transformation. Since pVI115 cannot replicate in these *E. coli* strains, the resulting Cm<sup>r</sup> transformants were considered pVI115 chromosomal integrants. The integration efficiency obtained with pVI115 in the absence of T $\text{paseIS}_{1223}$  was very low ( $\sim 10^{-8}$ ) compared to that of the strain carrying a T $\text{paseIS}_{1223}$  without a cloned promoter ( $\sim 10^{-6}$ ), as well as that obtained with T $\text{paseIS}_{1223}$  fused to the  $P_{\text{hlbA}}$  promoter (up to  $10^{-3}$ ). In all strains, the integration efficiency of pVI119 was very close to the background level observed in the absence of identified promoter ( $10^{-8}$  to  $10^{-7}$ ). These results clearly show that T $\text{paseIS}_{1223}$  triggers pVI115 integration using the  $P_{\text{junc}}$  substrate *in trans* and that the  $P_{\text{hlbA}}$  promoter drastically enhances the expression of T $\text{paseIS}_{1223}$  in *E. coli*. The transposon as developed in this work mimics the double-strand DNA intermediate and integrates as a nonreplicative element in the target sequence (Fig. 1C). The target sites of 19 integrants were determined by direct sequencing of genomic DNA (GATC Biotech) using the primer OLB215 (Table 2), which targets one transposon extremity (Fig. 1C). Fourteen insertions had occurred in different putative open reading frames; three were located in noncoding regions and two in repetitive extragenic palindromic (REP) sequences (data not shown). To confirm ran-

Received 20 February 2012 Accepted 3 May 2012

Published ahead of print 18 May 2012

Address correspondence to Pascale Serror, pascale.serror@jouy.inra.fr, or Jean-François Cavin, cavinjf@u-bourgogne.fr.

\* Present address: Sophie Brinster, Institut Cochin, Université Paris Descartes, CNRS (UMR 8104), Paris, France.

H.L.-S. and S.B. contributed equally to this work.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.00531-12

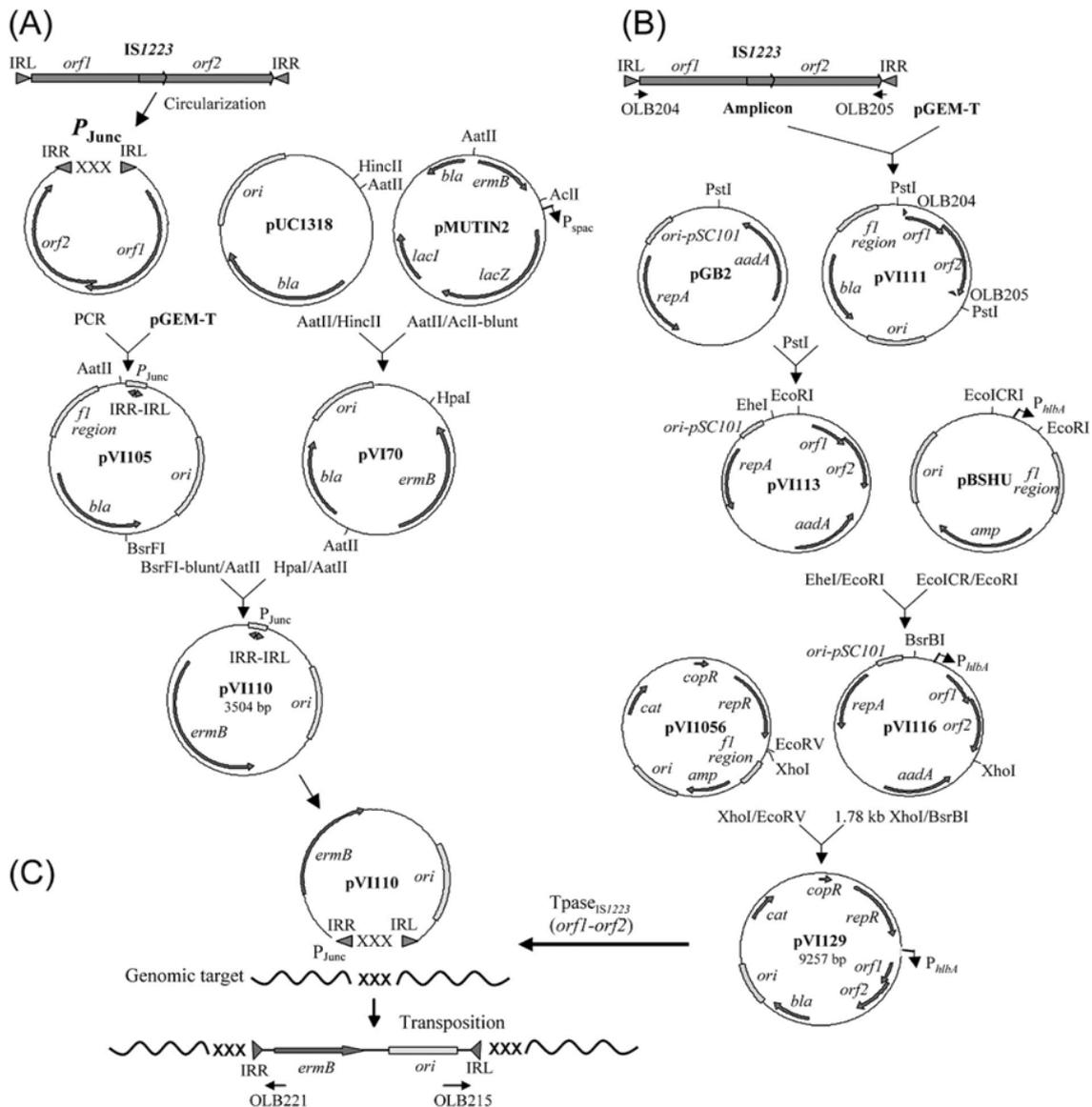


FIG 1 Diagram of plasmid construction. (A) Construction of the suicide transposon vector pV110. (B) Construction of the transposase IS1223 (*orf1-orf2*)-delivering vector pV1129. *orf1-orf2* are expressed under the promoter of the *L. delbrueckii* subsp. *bulgaricus* *hblA* gene ( $P_{hblA}$ ). The characteristics of each plasmid are indicated in Table 1. The genes *bla*, *aadA*, *cat*, and *ermB* encode resistance to ampicillin, spectinomycin, chloramphenicol, and erythromycin, respectively. (C) Map of integration of pV110 into genomic DNA by the action of T<sub>pase</sub><sup>IS1223</sup> on  $P_{junc}$ , with indication of primers OLB221 and OLB215 (Table 2) used for sequencing. XXX corresponds to the 3 to 4 base pairs duplicated during integration of pV110 in the genomic target. Plasmids are not drawn to scale.

domness of integration and saturation of the chromosome by pV1115, a pV115-mutagenized Lac<sup>+</sup> *E. coli* strain culture was diluted and spread onto LB medium with X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside) to screen Lac<sup>-</sup> mutants. Of the 10 Lac<sup>-</sup> clones analyzed, none was redundant, strongly supporting the randomness of integration and saturation of the *E. coli* chromosome by pV1115. Altogether, these results validated the fact that T<sub>pase</sub><sup>IS1223</sup> is active in *E. coli* and efficiently recognizes  $P_{junc}$  as a substrate leading to random transposition.

$P_{junc}$ -T<sub>pase</sub><sup>IS1223</sup> *in vivo* transposon mutagenesis in *L. delbrueckii* subsp. *bulgaricus* and in *L. casei*. The highly efficient  $P_{junc}$ -T<sub>pase</sub><sup>IS1223</sup> transposition system was adapted to LAB, namely, *L. delbrueckii* subsp. *bulgaricus* and *L. casei*.  $P_{hblA}$ -T<sub>pase</sub><sup>IS1223</sup> was cloned in the *E. coli*-Gram-positive bacterium shuttle vector, pV11056, to give pV1129 (Fig. 1B and Table 1), a plasmid providing the T<sub>pase</sub><sup>IS1223</sup>. Plasmid pV1129 possesses the pIP501 replication origin, which is thermosensitive in several Gram-positive bacteria (4, 15, 21, 26, 30), including *L. delbrueckii* subsp. *bulgaricus* (31). This property allows the efficient elimination of the

TABLE 1 Bacterial strains and plasmids

Strain or plasmid	Relevant markers, phenotypes, characteristics, and construction	Reference or source
<b>Strains</b>		
<i>E. coli</i>		
TG1	<i>supE hsdΔ5 thi Δ(lac-proAB) F' [traD36 proAB<sup>+</sup> lac<sup>n</sup> lacZΔM15]</i>	16
TG1repA	TG1 derivative with <i>repA</i> gene integrated into the chromosome	18
TG1pGB2	TG1 plus pGB2	This work
TG1pV1116	TG1 plus pV1116	This work
TG1pV1113	TG1 plus pV1113	This work
<i>L. delbrueckii</i> subsp. <i>bulgaricus</i>		
V1104	ATCC 11842 type strain	32
LBpV1129	V1104 carrying pV1129	This work
LBpV11056	V1104 carrying pV11056	This work
<i>L. casei</i>		
LC334	ATCC 334 type strain	Collection Institut Pasteur, France
LCpV1129	<i>L. casei</i> ATCC 334 carrying pV1129	This work
LCpV11056	<i>L. casei</i> ATCC 334 carrying pV11056	This work
<b>Plasmids</b>		
For construction of pV1115 and pV1116 and experiments with <i>E. coli</i>		
pBluescriptSK <sup>-</sup>	Ap <sup>r</sup> , pBR322ori	1
pBSHU	Ap <sup>r</sup> , pBluescriptSK <sup>-</sup> containing 317 bp of <i>L. delbrueckii</i> subsp. <i>bulgaricus</i> <i>hlyA</i> promoter region	9
pGB2	Sp <sup>r</sup> , pSC101ori	10
pGEM-T	Ap <sup>r</sup> , pBR322ori, <i>flori</i> , linear T-overhang vector	Promega
pJIM2242	Erm <sup>r</sup> , pWV01ori	18
pV142	Ap <sup>r</sup> , pBluescriptSK <sup>-</sup> IS1223 cloned at ClaI and EcoRI sites	31
pV162	pV142 with abutted P <sub>hlyA</sub> generated by cloning the two complementary oligonucleotides OLB187 and OLB188 between the ClaI site treated with exonuclease VII and the XhoI site of pV142	This work
pV1105	Ap <sup>r</sup> , pGEM-T with a 136-bp sequence containing P <sub>hlyA</sub> amplified with OLB131 and OLB203 primers using pV162 as a template	This work
pV1107	Cm <sup>r</sup> , pGEM-T containing the <i>cat</i> gene from pACYC184	This work
pV1108	Em <sup>r</sup> , pJIM2242 containing the P <sub>hlyA</sub> SpHI-PstI fragment (185 bp) of pV1105	This work
pV1111	Ap <sup>r</sup> , pGEM-T containing IS1223ΔIR	This work
pV1113	Sp <sup>r</sup> , pGB2 containing IS1223ΔIR	This work
pV1115	Cm <sup>r</sup> , pWV01oriΔ <i>repA</i> , P <sub>hlyA</sub> , obtained by ligation of the SpHI-EcoRII (extremity filled in with the Klenow fragment) P <sub>hlyA</sub> -containing fragment of pV1108 and the SpHI-HindI fragment (carrying the <i>cat</i> gene from pACYC184) of pV1107	This work
pV1116	Sp <sup>r</sup> , pGB2 containing P <sub>hlyA</sub> -IS1223ΔIR	This work
pV1119	Cm <sup>r</sup> , pV1115ΔP <sub>hlyA</sub> , obtained by self-ligating pV1115 digested with HincII-SchI	This work
pV1138	Em <sup>r</sup> , <i>E. coli</i> - <i>L. delbrueckii</i> subsp. <i>bulgaricus</i> shuttle vector	This work
For construction of pV1110 and pV1129 and experiments with lactobacilli		
pGB3631	Em <sup>r</sup> , pIP501 derivative	6
pGKV259	Cm <sup>r</sup> , Gram <sup>-</sup> /Gram <sup>+</sup> shuttle vector	38
pIP501	Em <sup>r</sup> , Gram <sup>-</sup> /Gram <sup>+</sup> shuttle vector, containing the replication origins from pBluescriptSK <sup>-</sup> for Gram <sup>-</sup> and from pIP501 for Gram <sup>+</sup> , including the copy number-controlling <i>copR</i> gene	22
pMUTIN2	Ap <sup>r</sup> Em <sup>r</sup> , pBR322ori	37
pUC1318	Ap <sup>r</sup> , pBR322ori	24
pV170	Em <sup>r</sup> , pUC1318 containing the <i>ermB</i> gene from pMUTIN2	This work
pV1110	Em <sup>r</sup> , pBR322ori, P <sub>hlyA</sub>	This work
pV1129	Ap <sup>r</sup> Cm <sup>r</sup> , pV11056 containing P <sub>hlyA</sub> -IS1223ΔIR	This work
pV1137	Em <sup>r</sup> , pBR322ori	This work
pV11052	Ap <sup>r</sup> Em <sup>r</sup> , pBR322ori, pIP501ori, obtained by ligation of pGB3631 and pBluescriptSK at EcoRI-BamHI sites	This work
pV11056	Ap <sup>r</sup> Cm <sup>r</sup> , pBR322ori, pIP501ori, <i>cop<sup>+</sup></i> , low-copy-number replicative plasmid in <i>L. delbrueckii</i> subsp. <i>bulgaricus</i> , obtained by ligation of the XhoI fragment (extremities filled in with the Klenow fragment) containing the Cm resistance ( <i>cat-86</i> ) cassette fragment of pGKV259 and the Eco47III-XbaI (extremity filled in with the Klenow fragment) of pV11052	This work

transposase-expressing plasmid so as to avoid further transposition events. The P<sub>hlyA</sub> sequence was combined with the erythromycin resistance cassette *ermB* to generate pV1110 (Fig. 1A), the suicide transposon plasmid.

*L. delbrueckii* subsp. *bulgaricus* V1104 and *L. casei* LC334 cells were first transformed with pV1129 as previously described (2, 32), and the resulting strains, LBpV1129 and LCpV1129, respectively, were then electroporated with 4 μg of pV1110, an optimal amount determined by preliminary assays with different amounts of plasmid DNA (data not shown). The cells were directly plated on MRS agar plates supplemented with erythromycin (5 μg/ml),

and the plates were incubated for 2 days at 42°C or 37°C for V1104 and LC334 strains, respectively, under static anaerobic growth conditions. The Em<sup>r</sup> colonies obtained after transformation with suicide transposon pV1110 were considered genomic (chromosome or indigenous plasmid) integrants. The transposition efficiency was determined by the number of Em<sup>r</sup> colonies obtained for 50 μl of electrocompetent cells with 1 μg of pV1110 plasmid. With LBpV1129, the number of integrants was estimated between 300 and 1,500 for ~2 × 10<sup>8</sup> viable cells, while with LCpV1129, this number reached between 2,500 and 7,500 for ~10<sup>9</sup> viable cells in more than 10 independent experiments. These results demon-

Downloaded from http://aem.asm.org/ on September 7, 2012 by INSTITUT PASTEUR-Médiathèque

TABLE 2 Primers

Primer use and name	Sequence (5' to 3') <sup>a</sup>	Target
<b>Plasmid construction</b>		
ERYF	GTTGATAGTGCAGTATCTTA	<i>ermC</i>
ERYR	CTTGCTCATAAGTAACGGTAC	<i>ermC</i>
IRLL-P <sub>junc</sub>	GATCTTTATGTCTAACAAATTATGAGGC	pVI62
IRLR-P <sub>junc</sub>	AAGTGCCTCATAAATTGTTAGACATAAAACGACTCCTGTAATAACAG	pVI62
M13rev	AACAGCTATGACCATG	pVI62
OLB93	AATGTAGGAAAGAAAGCACC	pVI62
OLB131	ACGACTCCTGTAATAACAG	M13-OLB93 amplicon
OLB187	TCGAATGTCTAACTTTTCTATGGCACTTC	Complementary to OLB188
OLB188	GAAGTCCATAGAAAAGTTAGACAT	Complementary to OLB187
OLB203	AAAATACCTCATAATTATTAGATTTTATGTCTAACAAATTATGAGGCAC	M13-OLB93 amplicon
OLB204	AAATCTGCAGTTATGAGGTATTTTTTATGACC	IS1223
OLB205	ACATTTCTCGAGTTTTAAAGATTTGATAATACACG	IS1223
<b>pVII10 target sequencing</b>		
OLB215	ATGGCCGCGGGATTACGACTCC	pVII10
OLB221	AGCTATGCATCCAACGCGTTGGG	pVII10
<b>Plasmid copy no. in <i>L. casei</i></b>		
LSEI0004F	ACCACCACAAGTTTGAAGG	LSEI_0004
LSEI0004R	TCACGCTCTTGCTAATGTCC	LSEI_0004
LSEI0145F	CGAAACCGAGGACTTGTG	LSEI_0145
LSEI0145R	AATGTGCGGGCTGAGAAC	LSEI_0145
LSEIA04F	ACTGGCACCAACGGATAGTC	LSEA_04 (pLSEIA)
LSEIA04R	GATGGCATTGAGACGACAGA	LSEA_04 (pLSEIA)
LSEIA13F	TTTGTTCGCTATCGGTTTCC	LSEIA_13 (pLSEIA)
LSEIA13R	AGTGGTTGATCGCACGACTA	LSEIA_13 (pLSEIA)

<sup>a</sup> Underlined bases indicate PstI restriction sites.

strate that P<sub>junc</sub>-T<sub>pase</sub>IS<sub>1223</sub> is functional and efficient in the two species. Negative controls were made using two other combinations of plasmids introduced successively: (i) pVII1056, as a T<sub>pase</sub>IS<sub>1223</sub>-nonexpressing vector, and pVII10 and (ii) pVII129 and pVII137, a P<sub>junc</sub>-less plasmid. For the two strains, less than 10 Em<sup>f</sup> colonies were obtained using these plasmid combinations.

These last data show that the P<sub>junc</sub> is not the substrate of a genomic indigenous putative transposase produced by *L. delbrueckii* subsp. *bulgaricus* or *L. casei* and that P<sub>junc</sub> is strictly required for pVII10 integration in these two species.

Preliminary results (data not shown) revealed that the growth of *L. casei* was seriously affected at temperatures above 40°C, mak-

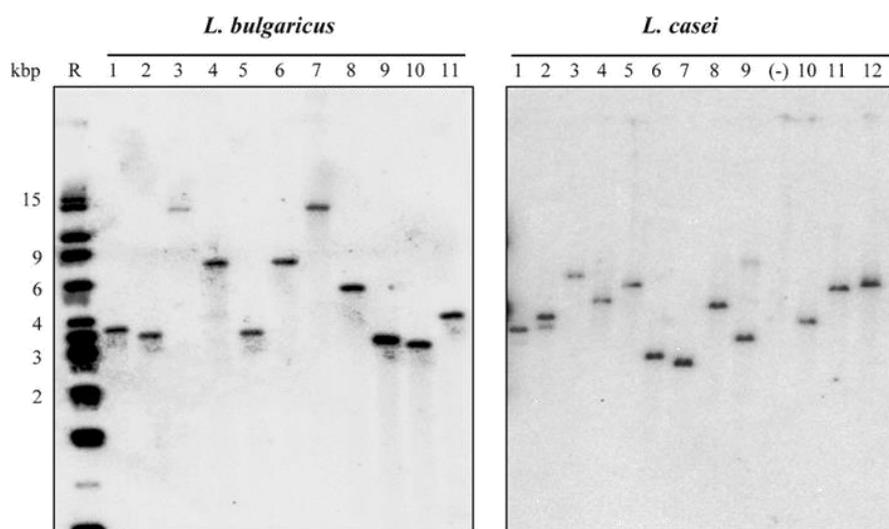


FIG 2 Diversity of pVII10 integration in *L. delbrueckii* subsp. *bulgaricus* and *L. casei*. Southern analysis of 11 *L. delbrueckii* subsp. *bulgaricus* integrants' DNA restricted by *NogMIV* and 12 *L. casei* integrants' DNA restricted by *HindIII* with a <sup>32</sup>P-labeled pVII10-*ermB* probe. R, Raoul molecular weight marker; (-), no DNA sample.

TABLE 3 Identification of target in independent *L. delbrueckii* subsp. *bulgaricus* and *L. casei* pVI110 integrants

Species and integrant <sup>a</sup>	pVI110 insertion site sequence <sup>b</sup>	Locus of pVI110 insertion <sup>c</sup>
<i>L. delbrueckii</i> subsp. <i>bulgaricus</i>		
Lb1	TTTCTTGGAAATTAAGCGCATAGTTA AATCACTTCTTTCTTTTCTTCTTTCTT	IGR Ldb2182/Ldb2183
Lb2	TAAAAAGTCTCGCTGAAAAGCGGGA CTTTTTGCGCCTTTGACGTGATTTTACA	IGR Ldb1406/Ldb1407
Lb3	AGATCATTCTCAAAAAGAGCTCCCG GAATCCGGGAGCTCTTTTGCTTTAGTTAA	Ldb0913
Lb4	AAAAAATACGAAGCAAAAGCAAGAA GAAAAAGCATGTTTGAAAAACATGCTTT	IGR Ldb0494/Ldb0501
Lb5	TGAGACCTATGTAAGAAGCTCAGGTC CACAGGACCGAGCTTCTTTTGCTTTTA	IGR Ldb2110/Ldb2112
Lb6	CATAAGCAACAAAAGCAGTCATTC CATCGATGACTGCTTTTCTGCTGCTGTGT	IGR Ldb0218/Ldb0219
Lb7	CGATAAAAAGAAAGGTCAGCGCGGCA AAAAGCGCTGGCCTTTTAAATTAGATT	IGR Ldb2015/Ldb2020
Lb8	CAGAATTTAGAGCAAAAGTAAAAGCCG CTTTTACCGGCTTTTATTTTTTCTTGT	IGR Ldb1491/Ldb1492
Lb9	GATGACAAAACAGGCTGAGGCCTAT TTTTATTTGCTTTTCTTCTTTTCTT	IGR Ldb1733/Ldb1734
Lb10	CAATAGCAAAAGAAAACTAGCTGA AAAACTAGCTAGTTTCTTTTCTTTCCCGT	IGR Ldb0968/Ldb0968
Lb11	ACAAAGCTTAAAAAGCGCTACAGGA CAACTTGCAGCGCTTTTATGTTTGTAA	IGR Ldb0164/Ldb0165
Lb12	AAGCCGCAACTACGGAATCTTGGA CTTGCCAAAAGCCCGGTACTTTTTCCCG	Ldb2064
Lb13	ATGTAGAAAAGAAAAGAGCTGCTC AAAGTGAGTAGCTTCGTTTTTGTCTATTA	IGR Ldb2034/Ldb2036
Lb14	TGTAACCTAAACTAATCCTTTTGGCA ATTTTCTGGGCTTTTTTGTAAATTTT	Ldb1636
Lb15	ACAAAAATCTTGCTTAACCTAATTGCA TTATAAACGGCTTTTTTGAATTTTGTTA	IGR Ldb2086/Ldb2087
Lb16	ATACAAGGAAAAAAGAGCTCCAGAA CTTGCTAAAAGCTTCTGAAGCTCTTTCTT	IGR Ldb2086/Ldb2087
Lb17	AATCAAACGAAAAAGCTTCAGTAAAG CAATACTGAAGCTTTTTTTCATTGCTATTA	IGR Ldb2090/Ldb2091
Lb18	TAGCAATGAAAAAGCTTCAGTATTG CTTACTGAAGCTTTTTCGTTTGATTCTA	IGR Ldb2090/Ldb2091
Lb19	ATACAAGGAAAAAAGAGCTCCAGA ACTTGCTAAACGCTTCTGAAGCTCTTTCTT	IGR Ldb2086/Ldb2087
Lb20	AAGAAAGAGCTTCAGAAGCGTTAG CAAGTCTGGAGCTTTTTTCTTCTGTAT	IGR Ldb2086/Ldb2087
<i>L. casei</i>		
Lc1	CGCTGGCGGATTATGTGACACCGGAAA ATGACTGGGAGCCGCTCAATTTTTTCAG	LSEI_1278
Lc2	AAAAAAGCTCACGTTTTGCGACGTGAG CTTTTTTGGTGCCGTCAGAACAAAGTTA	IGR LSEI_1440/LSEI_1441
Lc3	AGTGAAGCTCCAGACCGTGAATACAC AACCGGTGAAAAAACCATCAACGGTCT	LSEI_1892
Lc4	GTCACCGATGACAGCGCCAAGCTTTTC CGCGATTGGCAAAAGATCAACCGTCT	LSEI_1979
Lc5	TACACTGATGTTGAGAGATCAACATCA GTGTACAGCTCTTTATTTTGGGCCA	IGR LSEI_2050/LSEI_2051
Lc6	TTTTTGGTTAAGGGCTTTAATTTAGC TTGTTTTCTAAGTACTTTGCGACAT	IGR LSEI_A13/LSEI_A14
Lc7	CITTTGTCTTATGCTGGGGATTGGAAT TCTTAGACTGTTTTTTTCGTTTTTTAC	LSEI_0106
Lc8	TAAAAAGTGGCCCGGTAATACTGC AACGAGGCCACTTTTTATATTTATGGG	IGR LSEI_2579/LSEI_2580
Lc9	ACTCAGGTGATTTACATAGCTCCATG TTGCTGAGAGCCTTTAATTTAGGCA	LSEI_0797
Lc10	TGACCGGCAGGGTCATTGTCGGAGCCA ACATAAATAGTGGCTGGCAATTGCCCT	LSEI_0548
Lc11	ATTCAAAAAAAGTAAAAAGACTTTGCT AAACACAATCCAGAAATTAAGGCAAAA	IGR LSEI_A13/LSEI_A14
Lc12	TGGCCCTGCGTAATTTGACTTGAACA ACTGTTGGAAAGTTCTTTAATTTTTCT	LSEI_0374
Lc13	GTTGGCAGTCAGCAAGTCGTTTAAAA GCAGTCACCAATCAGAAAGACTATGAT	LSEI_2769
Lc14	GACGAAAAAACAAGAAGGTA TACGCC TAAACGCCGGTACCTTCTTTATTATCT	IGR LSEI_0637/LSEI_0638
Lc15	ATCAAAGATACTAAACAGCTTCTTGA AGATTTTACAGCCTTCTAAACACCAT	IGR LSEI_0343/LSEI_0344
Lc16	TTCTTGCTCAACAAAAAACACCACG AGGGTGAAAAAGTTGGGGAACTTTT	IGR LSEI_2568/LSEI_2569
Lc17	TTCAGGTGACGAAAAACAGTTTACCG ATACGCAACTCGAACTGCTACGAGTT	LSEI_2855
Lc18	CTGAACTCTTTGGCTTGAAAAATCAG ATAGGTAGTTTGGACGTTCTATTTCCT	LSEI_1966
Lc19	CCATAAGGAACACATGCACAATGCCGA AAAAAGACCATTGCATTTGTGCGCCGA	IGR LSEI_1566/1567
Lc20	CGGCTTACTAAAAAGAAGCTATATCTG ATGCACAGCAATTCTGCTGGGCGCGATA	IGR LSEI_2333/LSEI_2334

<sup>a</sup> Lb and Lc indicate *L. delbrueckii* subsp. *bulgaricus* and *L. casei*, respectively, in integrant designations.

<sup>b</sup> Inverted repeat sequences are underlined, and vertical bars are pVI110 insertion sites.

<sup>c</sup> IGR, intergenic region.

ing the elimination of pVI129 at a high temperature undesirable. The segregational stability of pVI129 in *L. casei* at 37°C was estimated at 86% per generation as described by Heap et al. (20). Thus, the inherent pVI129 instability and the resulting loss of T<sub>pase</sub><sub>IS1,2,3</sub> in *L. casei* mutants considerably limit the risk of genomic instability of the mutants.

**Analysis of pVI110 integration in *L. delbrueckii* subsp. *bulgaricus* and *L. casei*.** pVI110 insertion mutants of *L. delbrueckii* subsp. *bulgaricus* and *L. casei* were randomly selected. Mutant genomic DNA digested by NgoMIV for *L. delbrueckii* subsp. *bulgaricus* and by HindIII for *L. casei* was analyzed by Southern hybridization with a pVI110-specific probe generated by PCR am-

plification with primers ERYF and ERYR (Table 2). Plasmid pVI110 integrated in a single locus in each mutant, except for mutant 2 of *L. casei*, which presented two bands of different intensity, suggesting two distinct mutants in the sample (Fig. 2). Overall, the diversity of fragment sizes among the tested clones indicated that pVI110 had inserted randomly into both the *L. delbrueckii* subsp. *bulgaricus* and *L. casei* genomes. *L. casei* strain LC334 carries pLSEIA (GenBank accession number NC\_008502.1), a 29-kbp indigenous plasmid. Since indigenous plasmids are often targets of preferential insertion, leading to a reduction in efficiency of random transposon mutagenesis in chromosomal targets (28), we determined the plasmid copy number (PCN) of

pLSEIA by quantitative PCR (qPCR). Real-time PCRs were performed as previously described (25) from whole DNA of *L. casei* with primer pairs LSEI0004F-LSEI0004R and LSEI0145F-LSEI0145R (for the chromosome) and LSEIA04F-LSEIA04R and LSEIA13F-LSEIA13R (for pLSEIA) (Table 2). The PCN was determined using the following equation:  $PCN = (Ec)^{C_T} / (Ep)^{C_{TP}}$ , considering different amplification efficiencies [ $E = 10^{(-1/\text{slope})}$ ] and cycle threshold ( $C_T$ ) values for the two amplicons (chromosome,  $c$ , and plasmid,  $p$ ) (33). The PCN of pLSEIA was estimated at  $2.8 \pm 1.4$  (mean  $\pm$  standard deviation) plasmid copies per chromosome (from 3 independent DNA extracts). Taking into account the respective sizes of pLSEIA (29 kbp) and the *L. casei* chromosome (2.9 Mbp), the theoretical percentage of pVII10 nonpreferential integration in pLSEIA should be between 1 and 5%.

To confirm the diversity of mutants and to identify the nature of the target sequences of the pVII10 transposon, the randomly selected mutants were identified by genomic DNA sequencing with primers OLB215 and OLB221 (Table 2), which target the transposon sequence extremities (Fig. 1C and Table 3). In regard to *L. delbrueckii* subsp. *bulgaricus*, more than 80% ( $n = 17$ ) of sequenced targets were located in intergenic regions (IGR). Although four mutants were obtained in the IGR Ldb2086/Ldb2087, the target sequences of the pVII10 insertions were different, suggesting that this region is most likely not a hot spot of integration. Noticeably, target sites are surrounded by inverted repeats predicted to form hairpins with  $\Delta G < -9$  Kcal (calculated with Oligo Analyser freeware). Alignment of pVII10 target sequences revealed a preferential insertion in A/T-rich regions, as seen for other mobile elements, like Tn1545 in *Clostridium* and *Listeria* (5, 8), and several insertion sequences (11, 28). The nucleotide sequences of pVII10-target junctions in *L. delbrueckii* subsp. *bulgaricus* also revealed a 3-bp (occasionally 4-bp) duplication generated upon integration. Analysis of the target sequences suggests that triplets C/A A/T T/A may be preferential target sites for pVII10 in *L. delbrueckii* subsp. *bulgaricus*. Of the 20 random pVII10 transposon targets sequenced for *L. casei* (Table 3), 50% ( $n = 10$ ) were located in intergenic regions, while the *L. casei* genome contains about 20% noncoding regions (7). Ten percent ( $n = 2$ ) of mutants correspond to two different integration sites of pVII10 in pLSEIA, which represents only twice the maximal theoretical rate. This result reveals that pLSEIA is not a significant preferential host for pVII10 integration, indicating that the presence of the pLSEIA plasmid in LC334 is not an obstacle to obtaining saturated mutagenesis libraries. In contrast to the results for *L. delbrueckii* subsp. *bulgaricus*, only 20% of the *L. casei* target sites were located in inverted repeats predicted to form hairpins. Moreover, no preferential insertion in A/T-rich regions was observed. Despite the general, presumed random insertion of most transposons, many of them show a target preference (for reviews, see references 11 and 28). This targeting could in fact be a result of selective means to avoid affecting host fitness and, consequently, to promote transposon dissemination. Since *L. delbrueckii* subsp. *bulgaricus* is closely related to *L. johnsonii*, the original host of IS1223 (39), IS1223 is likely to preferentially target noncoding sequences to preserve its host genome. Interestingly, this bias is reduced in *L. casei*, which is phylogenetically more distant from *L. johnsonii* (3, 40), and is reduced even further in *E. coli*, suggesting that insertion sequences may display a more random integration in phylogenetically distant bacteria.

In conclusion, this work describes the use of an IS3-like trans-

position mechanism to engineer a novel transposition system based on the  $P_{\text{junc}}\text{-T}p\text{aseIS}_{1223}$  two-plasmid system for Gram-positive bacteria. Our results demonstrate that this system is functional in *L. delbrueckii* subsp. *bulgaricus* and *L. casei* and produces a high rate of stable integrants (at least 10,000 mutants per transformation for *L. casei*) despite the relatively poor transformation rate of lactobacilli. This system presents the advantage of promoting transposition of a suicide plasmid which contains  $P_{\text{junc}}$  (pVII10) provided in *trans* with a helper plasmid (pVII129) supplying  $Tp\text{aseIS}_{1223}$ . Thanks to this design, no sibling clones from early transposition events (31) can appear, and as pVII10 is stably produced in *E. coli*, it can be easily manipulated by inserting a reporter gene or used for signature-tagged mutagenesis. In view of the efficient transposition activity observed in the species tested (e.g., *Bacillus subtilis*, *Lactococcus lactis*, *Lactobacillus plantarum*, and *Enterococcus faecalis*; unpublished results), this transposition system may have a broad application in Gram-positive bacteria, particularly in LAB.

#### ACKNOWLEDGMENTS

We thank P. Polard for stimulating discussions on transposition mechanisms which led us to initiate this work. We also thank S. Kulakauskas for sharing data on *L. lactis*. We are grateful to Ellen Arena for the English revision of the manuscript.

This work was supported by ERC Advanced grant HOMEOPITH.

#### REFERENCES

1. Altling-Mees MA, Sorge JA, Short JM. 1992. pBluescriptII: multifunctional cloning and mapping vectors. *Methods Enzymol.* 216:483–495.
2. Aukrust TW, Brurberg MB, Nes IF. 1995. Transformation of *Lactobacillus* by electroporation. *Methods Mol. Biol.* 47:201–208.
3. Berger B, et al. 2007. Similarity and differences in the *Lactobacillus acidophilus* group identified by polyphasic analysis and comparative genomics. *J. Bacteriol.* 189:1311–1321.
4. Bhowmik T, Fernandez L, Steele JL. 1993. Gene replacement in *Lactobacillus helveticus*. *J. Bacteriol.* 175:6341–6344.
5. Blouzard JC, Valette O, Tardif C, de Philip P. 2010. Random mutagenesis of *Clostridium cellulolyticum* by using a Tn1545 derivative. *Appl. Environ. Microbiol.* 76:4546–4549.
6. Brantl S, Kummer C, Behnke D. 1994. Complete nucleotide sequence of plasmid pGB3631, a derivative of the *Streptococcus agalactiae* plasmid pIP501. *Gene* 142:155–156.
7. Cai H, Thompson R, Budinich MF, Broadbent JR, Steele JL. 2009. Genome sequence and comparative genome analysis of *Lactobacillus casei*: insights into their niche-associated evolution. *Genome Biol. Evol.* 1:239–257.
8. Choi KH, Kim KJ. 2009. Applications of transposon-based gene delivery system in bacteria. *J. Microbiol. Biotechnol.* 19:217–228.
9. Chouayekh H, Serror P, Boudebbouze S, Maguin E. 2009. Highly efficient production of the staphylococcal nuclease reporter in *Lactobacillus bulgaricus* governed by the promoter of the *hlyA* gene. *FEMS Microbiol. Lett.* 293:232–239.
10. Churchward G, Belin D, Nagamine Y. 1984. A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. *Gene* 31:165–171.
11. Craig NL. 1997. Target site selection in transposition. *Annu. Rev. Biochem.* 66:437–474.
12. Dower WJ, Miller JF, Ragsdale CW. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* 16:6127–6145.
13. Duval-Valentin G, Marty-Cointin B, Chandler M. 2004. Requirement of IS911 replication before integration defines a new bacterial transposition pathway. *EMBO J.* 23:3897–3906.
14. Duval-Valentin G, Normand C, Khemic V, Marty B, Chandler M. 2001. Transient promoter formation: a new feedback mechanism for regulation of IS911 transposition. *EMBO J.* 20:5802–5811.
15. Evans RP, Jr, Macrina FL. 1983. Streptococcal R plasmid pIP501: endo-

- nuclease site map, resistance determinant location, and construction of novel derivatives. *J. Bacteriol.* 154:1347–1355.
16. Gibson TJ. 1984. Studies on the Epstein-Barr virus genome. Ph.D. thesis. Cambridge University, Cambridge, United Kingdom.
  17. Goryshin IY, Jendrisak J, Hoffman LM, Meis R, Reznikoff WS. 2000. Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes. *Nat. Biotechnol.* 18:97–100.
  18. Guedon E, Serror P, Ehrlich SD, Renault P, Delorme C. 2001. Pleiotropic transcriptional repressor CodY senses the intracellular pool of branched-chain amino acids in *Lactococcus lactis*. *Mol. Microbiol.* 40:1227–1239.
  19. Gury J, Barthelmebs L, Cavin J-F. 2004. Random transposon mutagenesis of *Lactobacillus plantarum* by using the pGh9:ISS1 vector to clone genes involved in the regulation of phenolic acid metabolism. *Arch. Microbiol.* 182:337–345.
  20. Heap JT, Pennington OJ, Cartman ST, Minton NP. 2009. A modular system for *Clostridium* shuttle plasmids. *J. Microbiol. Methods* 78:79–85.
  21. Hershfield V. 1979. Plasmids mediating multiple drug resistance in group B *Streptococcus*: transferability and molecular properties. *Plasmid* 2:137–149.
  22. Horodniceanu T, Bouanchaud DH, Bieth G, Chabbert YA. 1976. R plasmids in *Streptococcus agalactiae* (group B). *Antimicrob. Agents Chemother.* 10:795–801.
  23. Ito M, et al. 2010. A practical random mutagenesis system for probiotic *Lactobacillus casei* using Tn5 transposition complexes. *J. Appl. Microbiol.* 109:657–666.
  24. Kay R, McPherson J. 1987. Hybrid pUC vectors for addition of new restriction enzyme sites to the ends of DNA fragments. *Nucleic Acids Res.* 15:2778.
  25. Licandro-Seraut H, Gury J, Tran NP, Barthelmebs L, Cavin JF. 2008. Kinetics and intensity of the expression of genes involved in the stress response tightly induced by phenolic acids in *Lactobacillus plantarum*. *J. Mol. Microbiol. Biotechnol.* 14:41–47.
  26. Luchansky JB, Muriana PM, Klaenhammer TR. 1988. Application of electroporation for transfer of plasmid DNA to *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Listeria*, *Pediococcus*, *Bacillus*, *Staphylococcus*, *Enterococcus* and *Propionibacterium*. *Mol. Microbiol.* 2:637–646.
  27. Maguin E, Prévost H, Ehrlich SD, Gruss A. 1996. Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. *J. Bacteriol.* 178:931–935.
  28. Mahillon J, Chandler M. 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62:725–774.
  29. Marco ML, Pavan S, Kleerebezem M. 2006. Towards understanding molecular modes of probiotic action. *Curr. Opin. Biotechnol.* 17:204–210.
  30. Pujol C, Chedin F, Ehrlich SD, Janniere L. 1998. Inhibition of a naturally occurring rolling-circle replicon in derivatives of the theta-replicating plasmid pIP501. *Mol. Microbiol.* 29:709–718.
  31. Serror P, Ilami G, Chouayekh H, Ehrlich SD, Maguin E. 2003. Transposition in *Lactobacillus delbrueckii* subsp. *bulgaricus*: identification of two thermosensitive replicons and two functional insertion sequences. *Microbiology* 149:1503–1511.
  32. Serror P, Sasaki T, Ehrlich SD, Maguin E. 2002. Electrotransformation of *Lactobacillus delbrueckii* subsp. *bulgaricus* and *L. delbrueckii* subsp. *lactis* with various plasmids. *Appl. Environ. Microbiol.* 68:46–52.
  33. Skulj M, et al. 2008. Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. *Microb. Cell Fact.* 7:6. doi:10.1186/1475-2859-7-6.
  34. Steinmetz M, Richter R. 1994. Easy cloning of mini-Tn10 insertions from the *Bacillus subtilis* chromosome. *J. Bacteriol.* 176:1761–1763.
  35. Tien MT, et al. 2006. Anti-inflammatory effect of *Lactobacillus casei* on *Shigella*-infected human intestinal epithelial cells. *J. Immunol.* 176:1228–1237.
  36. Tran NP, et al. 2008. Phenolic acid-mediated regulation of the *padC* gene, encoding the phenolic acid decarboxylase of *Bacillus subtilis*. *J. Bacteriol.* 190:3213–3224.
  37. Vagner V, Dervyn E, Ehrlich SD. 1998. A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* 144(Pt 11):3097–3104.
  38. van der Vossen JM, van der Lelie D, Venema G. 1987. Isolation and characterization of *Streptococcus cremoris* Wg2-specific promoters. *Appl. Environ. Microbiol.* 53:2452–2457.
  39. Walker DC, Klaenhammer TR. 1994. Isolation of a novel IS3 group insertion element and construction of an integration vector for *Lactobacillus* spp. *J. Bacteriol.* 176:5330–5340.
  40. Zhang ZG, Ye ZQ, Yu L, Shi P. 2011. Phylogenomic reconstruction of lactic acid bacteria: an update. *BMC Evol. Biol.* 11:1.

## I.2. Identification des sites d'insertion du transposon des mutants de *L. casei*

La mutagenèse par transposition présentée dans l'article précédent a permis d'obtenir environ 9250 mutants. Afin d'utiliser au mieux cette banque, nous avons décidé de séquencer le site d'insertion du transposon des mutants pour identifier la séquence interrompue.

Plusieurs techniques ont été testées, par exemple, la technique du "plasmid rescue" (digestion de l'ADN du mutant par des enzymes de restriction et auto-ligation puis transformation du produit ligué chez *E. coli*), qui est couramment utilisée pour identifier les sites d'insertion pour un nombre restreint de mutant. Cette technique n'était pas envisageable pour un nombre aussi important de mutants car trop longue et parfois peu efficace. D'autres techniques basées sur la PCR ont préalablement été utilisées chez d'autres bactéries (Jacobs *et al.* 2003; Quach *et al.* 2009) mais dans notre cas, nécessitent l'extraction d'ADN des mutants et ne peuvent être directement réalisées sur colonies pour une bactérie difficile à lyser comme *L. casei*, de plus de nombreux faux-positifs sont obtenus avec ces techniques (Bijlsma *et al.* 2007; Tong *et al.* 2004). De nouvelles techniques permettent à présent d'identifier l'ensemble des sites d'insertions (HITS et TraDIS) d'une banque de mutant cependant il est ensuite fastidieux de récupérer un mutant défini.

La mutagenèse ciblée est difficile à réaliser chez *L. casei* ATCC 334 probablement à cause de la présence de son plasmide, créant ainsi des incompatibilités avec d'autres plasmides exogènes. L'identification individuelle des mutants obtenus par mutagenèse par transposition permet d'avoir à portée de main, un large panel de mutants qui peuvent être utilisés et criblés pour différentes applications et recherches.

### Principaux résultats de l'article 2.

L'article suivant présente la méthode qui a été développée pour identifier les sites d'insertion du transposon de la banque de mutants de *L. casei*. L'ADN de chaque mutant a été extrait en plaques 96 puits à partir de cultures en phase exponentielle en utilisant le kit NucleoSpin 96 Tissue (Macherey Nagel). Plusieurs modifications par rapport aux recommandations du fournisseur ont été réalisées afin d'utiliser ce kit pour *L. casei*. Les cellules ont été soumises à un tampon de lyse auquel a été ajouté du Triton X-100 (le triton permet d'induire l'autolyse de la bactérie, voir résultats obtenus dans l'article 4), les temps d'incubations ont été augmentés afin d'obtenir une lyse complète de la bactérie. Puis, le volume d'éthanol permettant de précipiter l'ADN a été augmenté pour diminuer la viscosité de la solution avant le passage sur colonne de silice. La procédure de séquençage a, elle aussi, été modifiée afin de réaliser le séquençage directement sur l'ADN total. La bétaine a été ajoutée au mix de séquençage composé de l'amorce spécifique du transposon et de l'ADN pour réduire les structures secondaires de l'ADN. Le nombre de cycle de réaction de séquençage a été augmenté pour avoir un signal plus important. Après réaction de séquençage, les produits sont

purifiés en utilisant le kit BigDye Xterminator purification (Applied Biosystems) au lieu de la précipitation à l'éthanol.

La méthode mise au point a permis d'identifier le site d'insertion pour 8053 mutants sur 9250 soit 87% des mutants, elle est fiable et peut être utilisée sur d'autres bactéries. Sur un total de 20, 18 mutants par transposition de *Lactobacillus pentosus* ont été identifiés avec cette méthode et ont permis de valider les résultats obtenus par plasmid rescue pour l'article présenté en Annexe auquel, j'ai pu participer lors de mon travail de thèse.



## Rapid 96-well plates DNA extraction and sequencing procedures to identify genome-wide transposon insertion sites in a difficult to lyse bacterium: *Lactobacillus casei*



Hélène Scornec<sup>a</sup>, Magali Tichit<sup>b</sup>, Christiane Bouchier<sup>b</sup>, Thierry Pédrón<sup>c,d</sup>, Jean-François Cavin<sup>a</sup>, Philippe J. Sansonetti<sup>c,d,e,\*</sup>, Hélène Licandro-Seraut<sup>a,c</sup>

<sup>a</sup> UMR A PAM, AgroSup Dijon/Université de Bourgogne, 1 Esplanade Erasme, 21000, Dijon, France

<sup>b</sup> Plate-forme Génomique, Département Génomes et Génétique, Institut Pasteur, 28 Rue du Docteur Roux, 75724, Paris cedex 15, France

<sup>c</sup> Unité de Pathogénie Microbienne Moléculaire, Institut Pasteur, 28 Rue du Docteur Roux, 75724, Paris cedex 15, France

<sup>d</sup> Unité INSERM 786, Institut Pasteur, Paris, France

<sup>e</sup> Chaire de Microbiologie et Maladies Infectieuses, Collège de France, 11 Place Marcelin Berthelot, 75005, Paris, France

### ARTICLE INFO

#### Article history:

Received 16 June 2014

Received in revised form 4 August 2014

Accepted 5 August 2014

Available online 16 August 2014

#### Keywords:

Lactic acid bacteria  
Lactobacillus casei  
transposon mutagenesis  
Sanger sequencing  
bacterial genetics  
DNA extraction

### ABSTRACT

Random transposon mutagenesis followed by adequate screening methods is an unavoidable procedure to characterize genetics of bacterial adaptation to environmental changes. We have recently constructed a mutant library of *Lactobacillus casei* and we aimed to fully annotate it. However, we have observed that, for *L. casei* which is a difficult to lyse bacterium, methods used to identify the transposon insertion site in a few mutants (transposon rescue by restriction and recircularization or PCR-based methods) were not transposable for a larger number because they are too time-consuming and sometimes not reliable. Here, we describe a method for large-scale and reliable identification of transposon insertion sites in a *L. casei* mutant library of 9250 mutants. DNA extraction procedure based on silica membranes in 96-column format was optimized to obtain genomic DNA from a large number of mutants. Then reliable direct genomic sequencing was improved to fit the obtained genomic DNA extracts. Using this procedure, readable and identifiable sequences were obtained for 87% of the *L. casei* mutants. This method extends the applications of a library of this type, reduces the number of insertions needed to be screened, and allows selection of specific mutants from an arrayed and stored mutant library. This method is applicable to any already existing mutant library (obtained by transposon or insertional mutagenesis) and could be useful for other bacterial species, especially for highly lysis-resistant bacteria species such as lactic acid bacteria.

© 2014 Elsevier B.V. All rights reserved.

### 1. Introduction

Over the last decades, random transposon mutagenesis, coupled with adequate screening methods in bacteria (i.e. reverse genetics) has allowed major advances in identification and analysis of genes encoding unknown biological functions, including interactions of pathogenic bacteria with their host cells and tissues (Burrack and Higgins, 2007; West et al., 2005). It encouraged researchers to develop specific mutagenesis tools dedicated to their bacterial species of interest (Cummins et al., 2013; Mahillon and Chandler, 1998; Marcisisin et al., 2013; Wilson et al., 2009). The method encompasses a series of steps: (i) developing and performing random mutagenesis of the bacterial genome, (ii) assembling the mutant library, (iii) identifying the mutants of interest with an appropriate screening method, and (iv) analyzing the function of the gene products.

\* Corresponding author.

E-mail address: [philippe.sansonetti@pasteur.fr](mailto:philippe.sansonetti@pasteur.fr) (P.J. Sansonetti).

A completely annotated library offers major strategic advantages to analyze any function of the microorganism on a genome-wide basis. *A priori*, it allows appreciating the diversity, density and unbiased characteristics of the library, and it provides the option to reassemble mutants according to common functional properties of the mutated genes. *A posteriori*, it allows the validation of the particular role of a mutated gene or operon by allowing the selection of additional individual mutants with insertions in genes encoding different but related functions (i.e. metabolic pathways, regulatory systems, etc) to be further tested.

The choice of the method to identify the genomic DNA transposon insertion site depends notably on the species or strain of microorganism (GC content of DNA, effectiveness of bacterial lysis) and the type of transposon. Some methods are based on modified PCR amplification such as inverse PCR (Ochman et al., 1988), panhandle PCR (Jones and Winistorfer, 1992), or adaptor ligation (Rosenthal and Jones, 1990). They are designed to compensate for the fact that sequences around the transposon insertion site are unknown. Some reports described

identification by direct sequencing on a large DNA preparation (Horecka and Jigami, 2000; Qimron et al., 2003). However, the most commonly used technique is plasmid rescue, which consists in digestion of genomic DNA (gDNA) using restriction enzymes and recircularization of the transposon locus by self-ligation (Perpetuini et al., 2013; Tran et al., 2008). These methods have been used for microorganisms at a scale of 10 to 20 mutants.

Thus, establishing a reliable routine method for large-scale transposon mutant identification would be essential for many global reverse genetics studies. In *Lactobacillus* spp., few global reverse genetics studies have been reported due to several difficulties. The genetic manipulation of one given *Lactobacillus* strain is unpredictable without a long phase of development of genetic tools, transformation and lysis procedures (Aukrust et al., 1995; de Vos, 1999) because of their particularly difficult-to-lyse cell wall, their DNA restriction-modification systems, the presence of endogenous plasmids, and their variable level of electrocompetence. We have constructed a genome-wide mutant library of *Lactobacillus casei* by random transposon mutagenesis using the P<sub>unc</sub>-T<sub>pas</sub>I<sub>223</sub> system (Licandro-Seraut et al., 2012). When determining the transposon insertion site of several mutants, we faced all the above-mentioned difficulties when using the existing techniques; hence we needed to reconsider our strategy and methods to accurately identify transposon insertion sites in all the mutants. In this study, we report an improved and reliable method that was designed to extract pure gDNA from 96-well plate cultures of *L. casei* in sufficient quantity for direct sequencing of transposon insertion sites.

## 2. Materials And Methods

### 2.1. Bacterial Strains And Growth Conditions

Transposon mutants of *Lactobacillus casei* ATCC 334 (CIP 107868, Collection Institut Pasteur) used in this study belong to the tagged-mutant library generated (Licandro-Seraut et al., 2014) using tagged-derivative vectors of the P<sub>unc</sub>-T<sub>pas</sub>I<sub>223</sub> system as previously described (Licandro-Seraut, Brinster, van de Guchte, Scornec, Maguin, Sansonetti, Cavin and Serror, 2012). *L. casei* mutants were grown statically at 37 °C in MRS (Difco) supplemented with erythromycin (5 µg.ml<sup>-1</sup>).

### 2.2. Genetic manipulations

All DNA manipulations were performed according to standard procedures (Sambrook et al., 1989). Plasmids were isolated using gel and PCR clean up kit (Macherey-Nagel). PCR was performed using 0.1 unit of Platinum®Taq DNA polymerase high fidelity (Invitrogen), according to the manufacturer's recommendations, in an automatic thermocycler (Bio-Rad).

### 2.3. 96-well DNA extraction

The optimized protocol for 96-well plates DNA extraction was based upon the NucleoSpin 96 Tissue kit (Macherey Nagel). Several modifications of the manufacturer recommendation were made (Fig. 1). Cultures of *L. casei* (2 ml) in the exponential phase of growth (approximately 0.6 Unit of OD<sub>600 nm</sub>) were centrifuged and the pellets were resuspended in 180 µl of lysis buffer composed of TE (Tris-HCl 20 mM, EDTA 2 mM pH 8), lysozyme (20 mg.ml<sup>-1</sup>) and triton X-100 (1% v/v). After 2 h of incubation at 37 °C, 25 µl proteinase K were added and samples were incubated 3 h at 56 °C to obtain complete lysis. At this stage, the samples were frozen for storage overnight. Lysates were incubated for 5 min at room temperature with 20 µl of RNase A (5 mg.ml<sup>-1</sup>) (Invitrogen). Then, 200 µl of buffer BQ1 were added and incubated 10 min at 70 °C. At last, 400 µl of ethanol 96% (v/v) were added and half of the volume of the obtained mix was loaded on the column. The column plates were placed on 0.5 ml plates (Corning) and centrifuged at 3220 g for

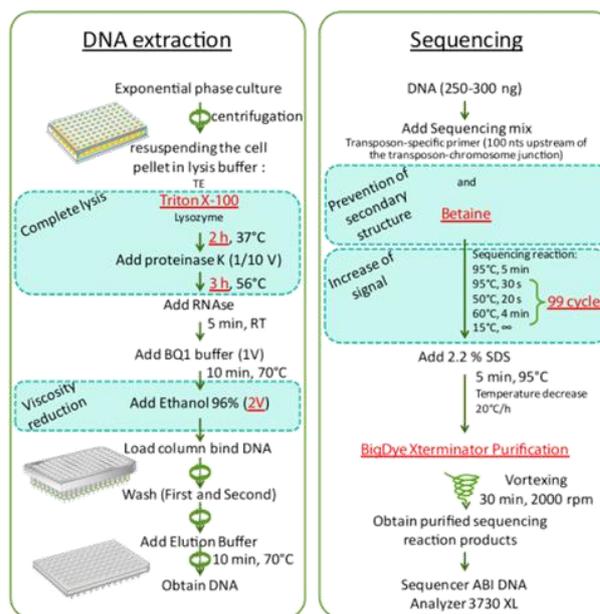


Fig. 1. Schematic representation of the DNA extraction and sequencing per 96-well plates of mutants. The main protocol optimization are in red letters.

6 min (Eppendorf 5810). The second half of the mix was passed through the column. The membranes of the columns were washed twice, firstly with 500 µl of BW solution (centrifugation: 4 min) and secondly with 500 µl of B5 solution (centrifugation: 5 min). To remove all traces of ethanol, the empty columns were centrifuged 2 min and heated 10 min at 70 °C. DNA was then eluted with 60 µl of elution buffer. DNA content was quantified by Qubit dsDNA HS kit (Invitrogen).

### 2.4. DNA sequencing

The insertion region was sequenced with the transposon-specific primer IRL (5'-GTTGGGAGCTCTCCCATATG-3') which anneals 100 nts upstream of the junction between the IRL extremity of the transposon and the *L. casei* genome because the sequencing chromatogram is generally not readable for the first nucleotides of the sequencing product. This primer meets the required specifications which are: a length between 18 and 22 nts, and a melting temperature between 55 and 60 °C. The reaction mix was composed of 4 µl of BigDye Terminator (V1.1) (Applied Biosystems) (Heiner et al., 1998), 2 µl of its corresponding BigDye sequencing buffer, 1 µl of primer (10 µM), 1 µl of betaine 5 M and 10 µl (250-300 ng) of DNA extract. Addition of betaine reduces dye signal, 99 cycles of sequencing reaction were performed as follow: a first cycle of denaturation at 95 °C for 5 min, 99 cycles of 95 °C, 30 s; 50 °C, 20 s; 60 °C, 4 min (Rusniok et al., 2010). The extension products were purified by addition of 2 µl of 2.2% SDS and heating at 95 °C for 5 min. After that, the temperature was slowly decreased to 20 °C/hour during 2 h. Then, the BigDye Xterminator purification kit (Applied Biosystems) was used and 20 µl of XTerminator™ Solution and 90 µl of SAM™ Solution were added in each sample. The purification was performed 30 min at 2000 rpm in MixMate device (Eppendorf). The purified sequenced products were analyzed on an ABI DNA Analyzer 3730 XL (Life technologies).

### 2.5. Determination of the transposon insertion sites

DNA sequences obtained were submitted against the genome of *L. casei* ATCC 334 and the nucleotide sequence of the transposon by

BLASTn (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>) (Altschul et al., 1990) to determine the transposon insertion site (Fig. 2). A threshold of  $10e^{-20}$  was applied for the e-value. The first nucleotide at the junction between the alignment of the transposon and the genome was considered as the nucleotide insertion.

### 2.6. Nucleotide sequence accession numbers

The nucleotide sequence of the transposon is in the NCBI nucleotide (<http://www.ncbi.nlm.nih.gov/nuccore>) under accession number U09558.1

## 3. Results And Discussion

### 3.1. Testing methods for large scale transposon target sequencing

Before designing and improving the method of direct genomic DNA sequencing described in this work, we sequentially tested several classical methods for identification of the genomic transposon insertion site in available mutants. An overview of the feasibility of these methods is presented in Table 1. The plasmid rescue method (Table 1, P1), which consists in digestion of genomic DNA using restriction enzymes, recircularization of the transposon locus by self-ligation and transformation of *E. coli*, appeared inefficient for large-scale mutant identification. This method is time-consuming and only fifty samples could be processed at once by one person especially because the *E. coli* transformation step requires the manipulation of mutant preparations on a one by one basis and its efficiency is uncertain. On a set of 14 mutants, we tried to bypass this transformation step by PCR-amplification of the genomic insert on the self-ligation product (Table 1, P2). However, despite various attempts to optimize the PCR-amplification step, multiple, non-specific amplicons were obtained for most of the mutants.

Among the methods used for transposon mutant identification, the PCR-based methods seemed most appropriate. In order to reduce the number of steps to a minimum, hence the cost and time of the method, we assessed three PCR-based methods using boiled colony as the matrix: the single prime PCR (Screaton et al., 1993), the random primer PCR (Lin et al., 1995) and the tailing 5' RACE PCR procedure (Leoni et al., 2008) (Table 1 P3). Despite the careful design of primers and the

optimization of PCR protocols, either non-specific amplicons or no amplicon were obtained except for one mutant. The correct insertion site identified corresponded to a *L. casei* plasmid locus. Indeed, this method preferentially purifies plasmid DNA than chromosomal DNA. We concluded that the amount and the purity of the DNA obtained after boiling the *L. casei* culture was insufficient to obtain specific PCR products and consequently a reliable mutant identification.

### 3.2. Identification of transposon insertion sites one by one using direct sequencing on genomic DNA

Considering the lack of reliability of undirect methods to identify the transposon insertion sites, we decided to focus on the development of a direct method based on direct genomic sequencing. Therefore, obtaining a pure gDNA preparation from *L. casei* became an essential step in order to produce readable and reliable sequences. This option reduced the number of steps, thus making it more time- and cost-effective, and decreasing the risks of non-specific amplifications.

The final procedure described in this work stems from several unsuccessful attempts to apply the above procedures to *L. casei*. Firstly, exponential-phase cultures (20 ml) of 20 randomly selected mutants were subjected to an appropriate enzymatic lysis (Lysozyme  $10 \mu\text{g}\cdot\text{ml}^{-1}$ , 3 h at  $37^\circ\text{C}$ ). Mutanolysin is often used in complement or to replace lysozyme to achieve lysis in *Lactobacillus* genus (Klaenhammer, 1984). In our hands, addition of 1000 U/ml of mutanolysin to the lysis buffer did not increase the rate of lysis. The obtained lysates were split in two aliquots in order to perform DNA extraction according to two parallel methods: (i) phenol-chloroform DNA extraction (Sambrook, Fritsch and Maniatis, 1989) (Table 1, P4) and silica column extraction (Nucleospin Tissue, Macherey-Nagel) according to the manufacturer's recommendations (Table 1, P5). The gDNA amount was  $20 \mu\text{g}$  for phenol-chloroform DNA extraction and  $2 \mu\text{g}$  for silica column extraction. The 40 genomic DNA preparations obtained were then subjected to direct sequencing. As transposon extremities are composed of inverted repeated sequences, betaine was added to the sequencing reaction mix to reduce the formation of secondary DNA structures. Knowing that the matrix corresponded to total gDNA, the number of cycles of sequencing was extended to 99 instead of 45 to 60 (for routine Sanger sequencing method using  $1 \mu\text{g}$  of gDNA). In the case of *L. casei* gDNA, the maximal gDNA amount in the

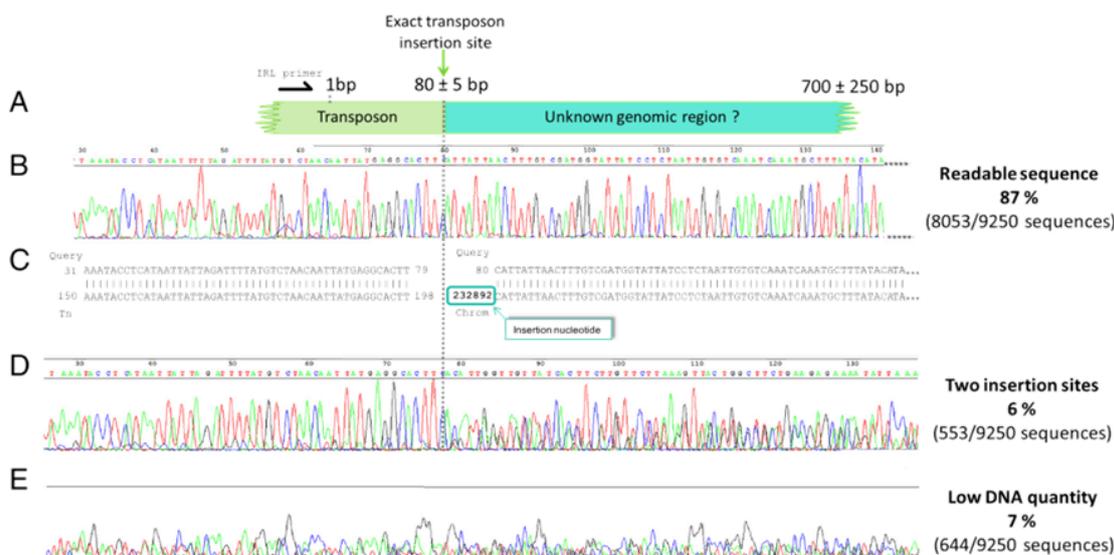


Fig. 2. Analysis of transposon insertion site in mutant library. A. Schematic representation of transposon-insertion site; B. Example of a sequencing chromatogram; C. Representative BLASTn against the *Lactobacillus casei* genome and transposon nucleotide sequence and identification of nucleotide insertion in the *L. casei* genome; D. Example of a sequencing chromatogram with transposon insertions in two different sites. E. Example of sequencing chromatogram with not enough DNA quantity.

**Table 1**  
Overview of feasibility and reliability of the tested procedures for transposon target identification in *L. casei*.

	Procedures	Time required for 1 person to process 768 (8x96) mutants	Mutants for which the right sequence was obtained
<b>P1</b>	<b>Plasmid rescue and transformation</b> <b>+ sequencing</b>	6 months * (by sets of 50 individual tubes)	35% (7 / 20 mutants)
<b>P2</b>	<b>Plasmid rescue and PCR</b> <b>+ sequencing</b>	6 months * (by sets of 50 individual tubes)	21% (3/14 mutants)
<b>P3</b>	<b>Single prime PCR</b> <b>+ sequencing</b>	3 days (using eight 96-well plates)	0% (1 / 20 mutants incl. 13 false positives)
<b>P4</b>	<b>DNA prep by Phenol/chloroform</b> <b>+ optimized sequencing</b>	2 months * (by sets of 50 individual tubes)	0% (0 / 20 mutants)
<b>P5</b>	<b>DNA prep by individual column</b> <b>+ optimized sequencing</b>	2 months * (by sets of 50 individual tubes)	30% (6 / 20 mutants)
<b>P6</b>	<b>Optimized DNA prep</b> <b>by 96-column plates</b> <b>+ optimized sequencing</b>	4 days (using eight 96-well plates)	87% (8053 / 9250 mutants)

\* Estimation

reaction mix was 300 ng. Typically, ethanol precipitation was used to purify sequencing reaction products; it consisted in precipitating DNA by ethanol, pelleting, drying and resuspending it. Here, to save a maximal quantity of sequencing reaction product, the BigDye Xterminator purification kit was chosen instead of ethanol precipitation before analysis on an ABI DNA Analyzer 3730 XL (Life technologies). Despite similar quality and quantity of gDNA preparations, no sequences were obtained with gDNA prepared by the phenol-chloroform procedure, probably due to remaining traces of phenol, whereas 6 sequences among the 20 gDNA preparations with silica column were obtained and proved to be correct following the sequencing of the PCR amplicons of the suspected transposon target sequences. The accuracy of sequencing results demonstrated the feasibility of this method which was subsequently optimized to increase the number of readable sequences, and adapted to a 96-well plate format.

### 3.3. Optimization of the identification of transposon insertion sites by set of 96 mutants using direct sequencing on genomic DNA

Individual cultures of *L. casei* mutants (200 µl) were grown for 36 hours. These 200 µl of cultures were mixed in 2 ml of fresh medium in deep-96-well plates, incubated for five hours (until the exponential phase of growth) and subjected to gDNA extraction as reported in the Materials and Methods section. Several modifications of the manufacturer's recommendation were made to adapt the protocol to lactic acid bacteria that are considered resistant to conventional lysis procedures (alkaline lysis), and to maximize the concentration of extracted gDNA (Fig. 1). In particular, individual silica columns were replaced by a 96-column device. Triton X-100 (1% v/v) was added to the lysis buffer because it induces autolysis and was essential to obtain a clear lysate. The clear lysate appeared extremely viscous and made passage through the silica membrane difficult, even applying a high speed centrifugation. To overcome centrifugation limitations, the lysate viscosity was reduced by doubling the volume of ethanol before the membrane-adsorption step. For each preparation, 1.5 to 2 µg of high quality double stranded gDNA were obtained with the 96-well plate method showing comparable efficiency in DNA extraction as the individual DNA column preparations, in spite of the fact that the volume of the culture was divided by five. Thanks to the optimization of each step, the gDNA preparation reached a higher concentration than the threshold for direct sequencing. Then each gDNA sample was subjected to sequencing as described in the Materials and Methods section and aligned by BLASTn against the transposon and the *L. casei* genome. The overall procedure to identify transposon mutants from a bacterial culture took 4 days in which 8 plates of 96 mutants were concurrently processed by one person.

Using this procedure, 9250 *L. casei* transposon mutants were processed and 8053 readable and identifiable sequences were obtained, corresponding to 87% of the processed mutants. The average length of the readable sequence was around 700 ± 250 nts (Fig. 2A, B and C), including 80 ± 5 nts of the transposon sequence. As the readings at the junction between the transposon and the gDNA were clear, it demonstrates the specificity of the direct sequencing procedure and enables a precise (at the nucleotide level) localization of the transposon insertion sites. Also, for twenty randomly selected mutants, transposon insertion sites were validated by PCR (data not shown). Unusable sequences correspond either to transposon insertions in two different sites, in this case, only the transposon sequence is readable (Fig. 2.D) or to signals under the detection level (Fig. 2.E). For these unusable sequences, a new extraction and sequencing could be done (i) to confirm transposon insertion in two different sites and absence of cross-contamination, then each insertion could be identified by plasmid rescue methods or (ii) to increase the concentration of gDNA.

## 4. Discussion

Due to the difficulty to determine the transposon insertion sites in mutants from random mutant library, some tools have been designed to facilitate this step (van Opijnen and Camilli, 2013). New sequencing methods like transposon sequencing (Tn-seq) (van Opijnen et al., 2009) and insertion sequencing (InSeq) (Goodman et al., 2009) were developed. However, they can only be used with a dedicated transposon (Himar I *Mariner* transposon). Other methods recently proved their efficiency regardless of the type of transposon like transposon-directed insertion-site sequencing (TraDIS) (Langridge et al., 2009) and high-throughput insertion tracking by deep sequencing (HITS) (Gawronski et al., 2009). Their objective is to quantify the variation of pooled mutants between two conditions. The procedure described here allows complete annotation of a transposon mutant library, with each mutant sequence individually identified and the mutants individually arrayed. Thus it is possible to go back to select a particular transposon mutant from the mutant library for further studies. Indeed the site directed mutagenesis is difficult to achieve in lactic acid bacteria. A thermosensitive plasmid is required, so several months may elapse before obtaining the desired mutant. Possessing an annotated library of mutants, and the ability to identify and use individual mutants from the original library, plus being able to reorganize that library according to the annotations of genome is of a great interest for any fundamental or applied researches on this species or any species.

To determine the transposon insertion site at the scale of a 9250 transposon mutant library, we have developed a reliable procedure based upon DNA preparations using silica membranes in 96-column format and genomic DNA direct sequencing. Our method fits well for

*L. casei*, a highly lysis resistant bacteria species, as most of the lactic acid bacteria (Klaenhammer, 1984). It allowed identifying almost 90% of transposon mutants with very high precision and reliability. This transposon insertion site determination method could be applied to other bacteria such as *Lactobacillus pentosus* for which 18 mutants over 20 were successfully identified (data not shown), and especially for other lactic acid bacterial species for which global interest has recently increased in the framework of functional (meta)genomic studies to characterize the role of the gut microbiota in health.

### Acknowledgements

This study was supported by the European Research Council Advanced grant HOMEOPITH to PJS, Grant agreement no: 232798. PJS is a Howard Hughes Medical Institute Foreign Scholar. HS was supported by a grant from the Ministère de l'Enseignement Supérieur et de la Recherche. We thank the Pasteur Institute Microorganism Collection for the *L. casei* Type strain (CIP 107868, ATCC 334).

### References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Aukrust, T.W., Brurberg, M.B., Nes, I.F., 1995. Transformation of *Lactobacillus* by electroporation. *Methods Mol. Biol.* 47, 201–208.
- Burack, L.S., Higgins, D.E., 2007. Genomic approaches to understanding bacterial virulence. *Curr. Opin. Microbiol.* 10, 4–9.
- Cummins, J., Casey, P.G., Joyce, S.A., Gahan, C.G., 2013. A Mariner transposon-based Signature-Tagged Mutagenesis system for the analysis of oral Infection by *Listeria monocytogenes*. *PLoS ONE* 8, e75437.
- de Vos, W.M., 1999. Gene expression systems for lactic acid bacteria. *Curr. Opin. Microbiol.* 2, 289–295.
- Gawronski, J.D., Wong, S.M., Giannoukos, G., Ward, D.V., Akerley, B.J., 2009. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proc. Natl. Acad. Sci. U. S. A.* 106, 16422–16427.
- Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., Gordon, J.I., 2009. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* 6, 279–289.
- Heiner, C.R., Hunkapiller, K.L., Chen, S.M., Glass, J.L., Chen, E.Y., 1998. Sequencing multimegabase-template DNA with BigDye terminator chemistry. *Genome Res.* 8, 557–561.
- Horecka, J., Jigami, Y., 2000. Identifying tagged transposon insertion sites in yeast by direct genomic sequencing. *Yeast* 16, 967–970.
- Jones, D.H., Winistorfer, S.C., 1992. Sequence specific generation of a DNA panhandle permits PCR amplification of unknown flanking DNA. *Nucleic Acids Res.* 20, 595–600.
- Klaenhammer, T.R., 1984. A general method for plasmid isolation in *Lactobacilli*. *Curr. Microbiol.* 10, 23–28.
- Langridge, G.C., Phan, M.D., Turner, D.J., Perkins, T.T., Parts, L., Haase, J., Charles, I., Maskell, D.J., Peters, S.E., Dougan, G., Wain, J., Parkhill, J., Turner, A.K., 2009. Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.* 19, 2308–2316.
- Leoni, C., Gallerani, R., Ceci, L.R., 2008. A genome walking strategy for the identification of eukaryotic nucleotide sequences adjacent to known regions. *Biotechniques* 44 (229), 225–232.
- Licandro-Seraut, H., Brinster, S., van de Guchte, M., Scornec, H., Maguin, E., Sansonetti, P., Cavin, J.F., Serror, P., 2012. Development of an Efficient *In Vivo* System ( $P_{\text{JusC}}\text{-TpaseIS}_{1223}$ ) for random transposon mutagenesis of *Lactobacillus casei*. *Appl. Environ. Microbiol.* 78, 5417–5423.
- Licandro-Seraut, H., Scornec, H., Pedron, T., Cavin, J.F., Sansonetti, P.J., 2014. Functional genomics of *Lactobacillus casei* establishment in the gut. *Proc. Natl. Acad. Sci. U. S. A.* 111, E3101–E3109.
- Lin, X., Kelemen, D.W., Miller, E.S., Shih, J.C., 1995. Nucleotide sequence and expression of *kerA*, the gene encoding a keratinolytic protease of *Bacillus licheniformis* PWD-1. *Appl. Environ. Microbiol.* 61, 1469–1474.
- Mahillon, J., Chandler, M., 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62, 725–774.
- Marcisain, R.A., Bartho, T., Bulach, D.M., Srikram, A., Sermswan, R.W., Adler, B., Murray, G. L., 2013. Use of a high-throughput screen to identify *Leptospira* mutants unable to colonize the carrier host or cause disease in the acute model of infection. *J. Med. Microbiol.* 62, 1601–1608.
- Ochman, H., Gerber, A.S., Hartl, D.L., 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* 120, 621–623.
- Perpetuini, G., Scornec, H., Tofalo, R., Serror, P., Schirone, M., Suzzi, G., Corsetti, A., Cavin, J. F., Licandro-Seraut, H., 2013. Identification of critical genes for growth in olive brine by transposon mutagenesis of *Lactobacillus pentosus* C11. *Appl. Environ. Microbiol.* 79, 4568–4575.
- Qimron, U., Madar, N., Ascarelli-Goell, R., Elgrably-Weiss, M., Altuvia, S., Porgador, A., 2003. Reliable determination of transposon insertion site in prokaryotes by direct sequencing. *J. Microbiol. Methods* 54, 137–140.
- Rosenthal, A., Jones, D.S., 1990. Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. *Nucleic Acids Res.* 18, 3095–3096.
- Rusniok, C., Couve, E., Da Cunha, V., El Gana, R., Zidane, N., Bouchier, C., Poyart, C., Leclercq, R., Trieu-Cuot, P., Glaser, P., 2010. Genome sequence of *Streptococcus gallolyticus*: insights into its adaptation to the bovine rumen and its ability to cause endocarditis. *J. Bacteriol.* 192, 2266–2276.
- Sambrook, J., Fritsch, E., Maniatis, T., 1989. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press.
- Screaton, G.R., Bangham, C.R., Bell, J.L., 1993. Direct sequencing of single primer PCR products: a rapid method to achieve short chromosomal walks. *Nucleic Acids Res.* 21, 2263–2264.
- Tran, N.P., Gury, J., Dartois, V., Nguyen, T.K., Seraut, H., Barthelmebs, L., Gervais, P., Cavin, J.F., 2008. Phenolic acid-mediated regulation of the *padC* gene, encoding the phenolic acid decarboxylase of *Bacillus subtilis*. *J. Bacteriol.* 190, 3213–3224.
- van Opijnen, T., Camilli, A., 2013. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat. Rev. Microbiol.* 11, 435–442.
- van Opijnen, T., Bodi, K.L., Camilli, A., 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* 6, 767–772.
- West, N.P., Sansonetti, P., Mounier, J., Exley, R.M., Parsot, C., Guadagnini, S., Prevost, M.C., Prochnicka-Chaloufour, A., Delepierre, M., Tanguy, M., Tang, C.M., 2005. Optimization of virulence functions through glucosylation of *Shigella* LPS. *Science* 307, 1313–1317.
- Wilson, A.C., Hoch, J.A., Perego, M., 2009. Two small c-type cytochromes affect virulence gene expression in *Bacillus anthracis*. *Mol. Microbiol.* 72, 109–123.

## II. Identification des gènes impliqués lors de l'établissement de *L. casei* dans l'intestin.

### Principaux résultats de l'article 3.

Le séquençage de la banque de mutant de *L. casei* présenté dans l'article précédent a permis d'identifier 3037 mutants ayant une insertion du transposon dans la séquence d'un gène (2787 intégrations du transposon dans les gènes du chromosome et 250 intégrations dans les gènes du plasmide). Les mutations sont réparties sur la totalité du génome comme peu l'attester la figure 1 de l'article suivant représentant la répartition des sites d'insertion du transposon et la banque de mutants représente une proportion de groupes fonctionnels similaire à la proportion du génome.

La banque de mutants de *L. casei* présentée dans l'article 1 a été réalisée par une mutagenèse aléatoire permettant de suivre les mutants : la Signature-Tagged Mutagenesis (STM) (Mazurkiewicz *et al.* 2006; West *et al.* 2005). En effet, la mutagenèse a été réalisée avec 70 plasmides pVII10 étiquetés différemment, c'est-à-dire que dans la séquence du transposon, une étiquette (tag) de 40 pb a été clonée, permettant ainsi de réaliser des cribles en ensemble (pool) de 70 mutants, ces mutants pouvant ensuite être identifiés grâce à leur étiquette. Le séquençage des sites d'insertions du transposon des mutants de la banque de *L. casei* a permis de réassembler la banque de mutants en une banque ne contenant que les mutants avec un gène interrompu en un seul exemplaire ; c'est-à-dire, que lorsque plusieurs mutants existaient pour un même gène, un seul a été utilisé pour le réassemblage. Au total, la banque réassemblée comporte 1110 mutants dans des gènes différents. La réorganisation a été réalisée tout en prenant en compte l'étiquette de chaque mutant afin de pouvoir réaliser les cribles en ensembles de 70 mutants. Dans le cas où l'assemblage n'était pas complet (moins de 70 étiquettes différentes), des mutants avec insertion du transposon dans les régions intergéniques ont été ajoutés afin d'obtenir au minimum 40 mutants par ensemble.

Nous voulions utiliser cette banque de mutants pour déterminer les gènes impliqués dans l'établissement de *L. casei* au niveau du tube digestif. Des essais de criblage ont, au préalable, été réalisés chez la souris par gavage gastro-intestinal mais les résultats n'étant pas reproductibles le modèle souris n'a pas été sélectionné pour la suite du criblage. L'anse iléale ligaturée de lapin, modèle déjà utilisé par l'équipe PMM (Marteyn *et al.* 2010; Sansonetti *et al.* 1983; West *et al.* 2005) a alors été choisi comme modèle.

Chaque ensemble de mutants a été inoculé dans des anses iléales ligaturées de deux lapins. Après 16 h de compétition *in vivo*, les anses sont prélevées, broyées et des dilutions sont étalées sur boîtes MRS afin de récupérer les mutants. La variation quantitative de chaque mutant, entre l'ensemble injecté et l'ensemble récupéré, a été réalisée par PCR quantitative (qPCR) grâce à l'étiquetage. En effet, 70 couples d'amorces spécifiques de chacun des 70 étiquettes ont été dessinés permettant ainsi de corrélérer la quantité de chaque mutant par la

quantité d'étiquettes. Ainsi 69 mutants ont été sélectionnés car leur quantité était diminuée *in vivo*. Leur croissance *in vitro* a été testée afin d'éliminer les mutants ayant un défaut de croissance par rapport à la souche parentale. Les mutants dans les gènes LSEI\_0281, 1278, 1468, 1565 et 1566 ont ainsi été éliminés. Les mutants restants ont été criblés de nouveau et l'ensemble injecté a cette fois-ci été étalé sur boîte afin d'éliminer les mutants qui avaient des difficultés à se développer. Ainsi 47 mutants ont été identifiés comme ayant des difficultés à persister *in vivo*. Les protéines codées par les gènes interrompus chez ces mutants ont des fonctions variées incluant, les processus cellulaires basiques, le métabolisme, la biosynthèse de la paroi cellulaire, l'adaptation au milieu environnant...

La banque annotée permet à la fois d'identifier les mutants, et par conséquent les gènes impliqués dans l'établissement de *L. casei* dans l'intestin mais aussi les gènes qui ne sont pas impliqués. La banque est donc doublement informative.



# Functional genomics of *Lactobacillus casei* establishment in the gut

Hélène Licandro-Seraut<sup>a,b,c</sup>, Hélène Scornec<sup>a</sup>, Thierry Pédrón<sup>b,c</sup>, Jean-François Cavin<sup>a</sup>, and Philippe J. Sansonetti<sup>b,c,d,1</sup>

<sup>a</sup>Unité Mixte de Recherche A 02102 Procédés Alimentaires et Microbiologiques, AgroSup Dijon/Université de Bourgogne, 21000 Dijon, France; <sup>b</sup>Unité de Pathogénie Microbienne Moléculaire and <sup>c</sup>Institut National de la Santé et de la Recherche Médicale U786, Institut Pasteur, 75724 Paris Cedex 15, France; and <sup>d</sup>Chaire de Microbiologie et Maladies Infectieuses, Collège de France, 75005 Paris, France

Contributed by Philippe J. Sansonetti, June 25, 2014 (sent for review April 1, 2014)

Although the composition of the gut microbiota and its symbiotic contribution to key host physiological functions are well established, little is known as yet about the bacterial factors that account for this symbiosis. We selected *Lactobacillus casei* as a model microorganism to proceed to genomewide identification of the functions required for a symbiont to establish colonization in the gut. As a result of our recent development of a transposon-mutagenesis tool that overcomes the barrier that had prevented *L. casei* random mutagenesis, we developed a signature-tagged mutagenesis approach combining whole-genome reverse genetics using a set of tagged transposons and in vivo screening using the rabbit ligated ileal loop model. After sequencing transposon insertion sites in 9,250 random mutants, we assembled a library of 1,110 independent mutants, all disrupted in a different gene, that provides a representative view of the *L. casei* genome. By determining the relative quantity of each of the 1,110 mutants before and after the in vivo challenge, we identified a core of 47 *L. casei* genes necessary for its establishment in the gut. They are involved in housekeeping functions, metabolism (sugar, amino acids), cell wall biogenesis, and adaptation to environment. Hence we provide what is, to our knowledge, the first global functional genomics analysis of *L. casei* symbiosis.

commensalism | Lactic acid bacteria

The pioneering studies that led to the characterization of the gut microbiota were reviewed in 2001 (1). These studies and recent investigations have revealed mutualistic functions (2), including a barrier effect against allogenic microbes (3), fermentation of complex sugars (4, 5), and maturation and homeostasis of the immune system (6). Recent metagenomic studies have revealed an extraordinary diversity of genes constituting the gut microbiome (7), opening the way to correlative studies linking microbiome diversity, homeostasis, and diseases (5, 8, 9).

In parallel, some representative species, i.e., “model symbionts,” now are being studied functionally (10). As it was done for pathogens, it is essential to develop the cellular microbiology of symbionts and particularly to identify the genes required for their establishment and persistence in the gut. Transcriptomic profiling identified up-regulated genes linked to metabolic functions, stress responses, and pili synthesis during early colonization (11–13). Comparative genomics among *Lactobacilli* identified strain-specific candidate genes for extended colonization: In *Lactobacillus rhamnosus*, persistence was attributed to an *spaCBA* locus encoding LPXTG-like pilins (14), and in *Lactobacillus johnsonii* it was attributed to specific glycosyltransferases, a phosphotransfer system, and a protease (15). Otherwise, a functional in vivo screening based on the expression of a genomic library of *Bacteroides fragilis* identified a locus encoding polysaccharide utilization as essential for stable colonization of murine colonic crypts (16). Alternatively, colonization of germ-free mice with a collection of random mutants of *Bacteroides thetaiotaomicron* followed by deep sequencing showed that mutants unable to synthesize vitamin B12 were impaired in gut colonization (17).

*Lactobacillus* spp. pioneer initial gut colonization (18), and they participate in the gut immunological and nutritional symbiosis. Because of our permanent exposure to *Lactobacillus* spp., and particularly to *Lactobacillus casei*, as components of dairy products, tools are required to decipher its symbiosis with the gut. Genetic manipulation of lactic bacteria often is problematic because of their natural resistance to numerous antibiotics and the lack of dedicated genetic tools and efficient transformation procedures (19). We recently developed a transposon-mutagenesis tool, the P<sub>junc</sub>-TpaseIS<sub>1223</sub> system (20), which overcomes the barrier to random mutagenesis in *Lactobacillus casei*. Using this tool, we adapted a signature-tagged mutagenesis (STM) approach (21) to *L. casei* that combined whole-genome reverse genetics using a set of tagged transposons with an in vivo screening in the rabbit ligated ileal loop model identifying mutants impaired in gut establishment. The term “establishment” qualifies the early steps of colonization explored by this model. After sequencing the 1,110 independent mutants obtained in this study, we identified a core of 47 *L. casei* genes belonging to five major functional groups that are required for its establishment in the gut.

## Results

**Generation of a Library of *L. casei* Tagged Mutants.** To generate a large library of *L. casei* tagged mutants and to proceed to STM, tagged derivatives of the P<sub>junc</sub>-TpaseIS<sub>1223</sub> transposable vector were generated using 70 DNA tags previously used for *Salmonella typhimurium* STM (21). For each tag, among the ~5,000 integrants obtained per transformation, clones were selected randomly and assembled in 96-well plates. Thus, a library of 9,250 tagged mutants labeled with 70 different tags was generated. To extend the contribution of STM, we introduced real-time PCR, rather than dot-blot analysis, to allow relative quantification of bacteria in addition to their detection.

## Significance

*Lactobacillus casei*, a food bacterium recognized for its beneficial effects, was selected as a model microorganism to proceed to genomewide identification of the functions required for a symbiont to establish colonization in the gut. We recently have developed a mutagenesis tool that overcomes the barrier that prevented *L. casei* random mutagenesis. After identifying 9,250 mutations, we assembled a library of 1,110 mutants disrupted in different genes and tested them for their ability to colonize an in vivo model, the rabbit ligated ileal loop. With this global functional genomic analysis of *L. casei* symbiosis (the first, to our knowledge), we identified a core of 47 *L. casei* genes necessary for its establishment in the gut.

Author contributions: H.L.-S., T.P., and P.J.S. designed research; H.L.-S., H.S., and T.P. performed research; H.L.-S., H.S., T.P., J.-F.C., and P.J.S. analyzed data; and H.L.-S., T.P., J.-F.C., and P.J.S. wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. Email: philippe.sansonetti@pasteur.fr.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1411883111/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1411883111/-DCSupplemental).

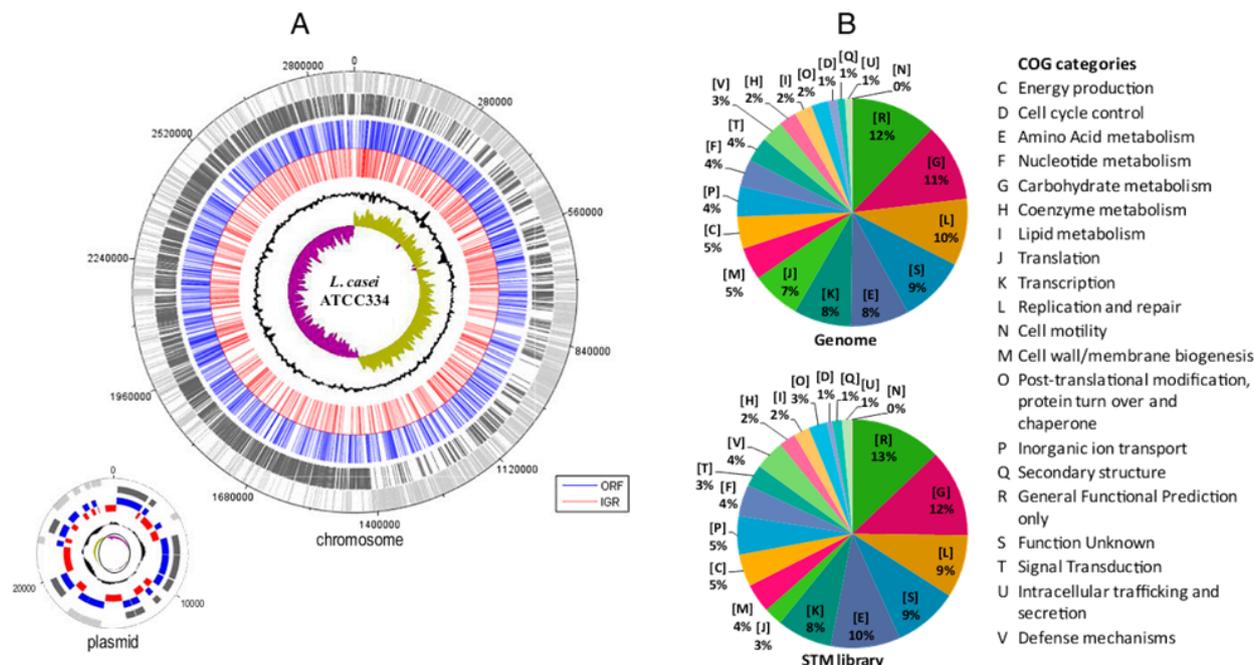
**Analysis and Assembly of a Library of *L. casei* Tagged Mutants.** Based on an initial screening showing that the intergenic regions contributed very little to gut establishment (*SI Text*), we focused on mutants in genes. Hence, we sequenced the transposon target in each mutant before in vivo screening (i.e., 9,250 mutants). Among the 8,053 readable sequences (87% of the total), 3,037 (37%) indicated that transposon integration occurred in a gene (2,787 integrations into the chromosome and 250 into the plasmid) (Fig. 1A). In most cases, when several mutations were found in the same gene, the sites of transposon insertion differed (84.5% of insertions in genes occurred in unique sites), and we selected one representative mutant. Mutations in intergenic regions (IGR) tended to occur in palindromic regions surrounding endogenous transposase genes (32.4% of the insertions in IGRs occurred in unique sites). In summary, our library of tagged *L. casei* mutants 1,096 mutants in distinct chromosomal genes and 14 mutants in distinct plasmid genes. Characterizing and assembling this library considerably reduced the number of mutants to be screened. The mutations appear to be evenly distributed throughout the genome (Fig. 1A). Overall, the proportion of the functional groups in the library is similar to their proportion in the genome (Fig. 1B). However, the proportion of genes encoding translation-related functions [with the COG letter (J)] is lower in the library, likely because of the high number of essential genes related to protein synthesis (22). Because most bacterial genes are organized in operons [1.5 genes per transcription unit predicted for *L. casei*, according to genome analysis using the Biocyc website (23)], transposon integration can cause a polar effect; thus the location of each gene of interest in a given operon is specified.

**Screening of Mutants in the Rabbit Ligated Ileal Loop Model and Identification of Crucial *L. casei* Genes for Establishment in the Gut.** The rabbit ligated ileal loop model allowed the screening of

a large number of *Shigella* mutants in competitive pools in the gut (24–26), and it explores the ileal conditions in which *L. casei* has been shown to persist in an active physiological state (27).

Twenty-two pools containing 25–70 differently tagged mutants were organized from the library of 1,110 discrete mutants (Fig. 2). The pools then were challenged in rabbit ligated ileal loops ( $5.10^7$  cfu per loop) for 16 h (Fig. 2). Each pool was tested twice in different rabbits. The ratio of each mutant to the whole pool was determined using tags for the injected and the recovered pools (Fig. 2). Based on the observed variation of mutant quantities in a pool, we considered a threefold difference in quantity between the injection and recovery pools a relevant threshold (*SI Text* and Fig. S1). Sixty-nine mutants were selected at this step because their quantity decreased at least threefold in two rabbits. We identified the mutants that were exclusively altered in vivo and did not show a growth defect in vitro by comparing the colony size and the turbidity of cultures at stationary phase in MRS medium of the mutants and the wild-type strain (Table S1). This comparison allowed us to eliminate six mutants (i.e., LSEI\_0281, 1278, 1468, 1565, and 1566). To eliminate growth biases linked to pool composition, the resulting 63 mutants were reassembled to create five new pools. Each pool was composed of 12 or 13 mutants being retested and 12 mutants picked randomly from the IGR library. A ratio of one mutant to 70 competitors was maintained. The in vivo challenge and the quantification methods were identical to those used in the first round, except that, to eliminate false positives caused by growth deficiencies on plates, the injected pool was quantified from colonies obtained by plating.

This scheme of selection identified 47 mutants of *L. casei* (4% of the screened genes) that were altered in their capacities of establishment in the gut (Table 1). The mutated genes encompassed various functions, including housekeeping, metabolism, biogenesis of the cell wall, and adaptation to environment. Most



**Fig. 1.** Analysis of transposon insertions in the genome of the 9250 *L. casei* mutants. (A) Genome atlas of transposon insertions. Circle 1: genes on the positive strand in light gray; circle 2: genes on the negative strand in dark gray; circle 3: disrupted genes in blue; circle 4: disrupted IGRs in red; circle 5: GC content; circle 6: GC skew. (B) Relative abundance of cluster of orthologous groups (COG) functional categories of genes in the *L. casei* genome and in the disrupted genes of the STM library. Hits found in COG represent 68.3% of the genome and 69% of the disrupted ORFs of the STM library.

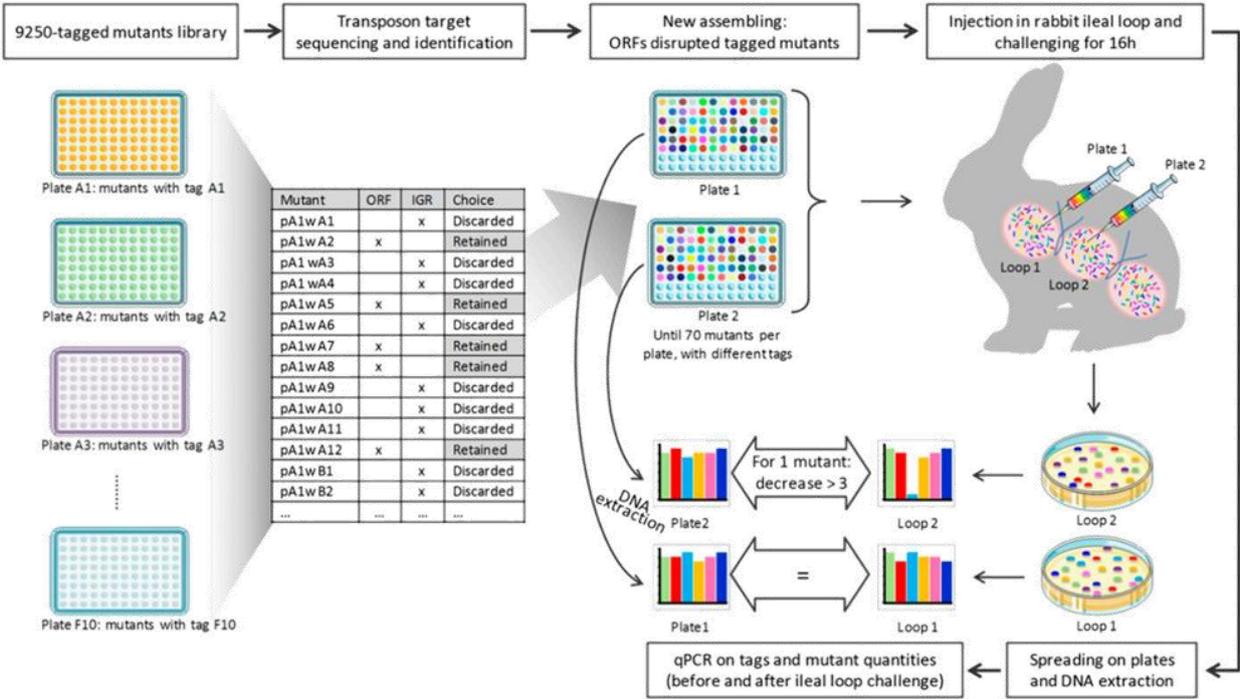


Fig. 2. Generation of the gene tagged-mutants library and its screening in rabbit ligated ileal loops.

are conserved among the Firmicute phylum. LSEI\_0238, 0242, and 0247 orthologs are only found in the *Lactobacillus* genus. Seven genes, LSEI\_0135, 0156, 0806, 1461, 1790, 2553, and 2601, are specific to *L. casei* or related species such as *L. rhamnosus*, *Lactobacillus paracasei*, or *Lactobacillus zeae*.

**Mutant Analysis. Housekeeping functions.** Our screening identified several housekeeping genes that are likely to help maintain

physiological parameters during the changes to which bacteria are suddenly exposed in the in vivo environment (Table 1 and Fig. 3): LSEI\_0757 and 1539 (nucleotide metabolism), LSEI\_1615 and 1488 (DNA replication), LSEI\_1668 (transcription), LSEI\_1656 (translation), and LSEI\_1274 (possibly implicated in cell division). We also identified the RecG family helicase RecG (LSEI\_1615) and DinG (LSEI\_1488), primarily involved in DNA replication and possibly in DNA repair (28). Mutations in this category

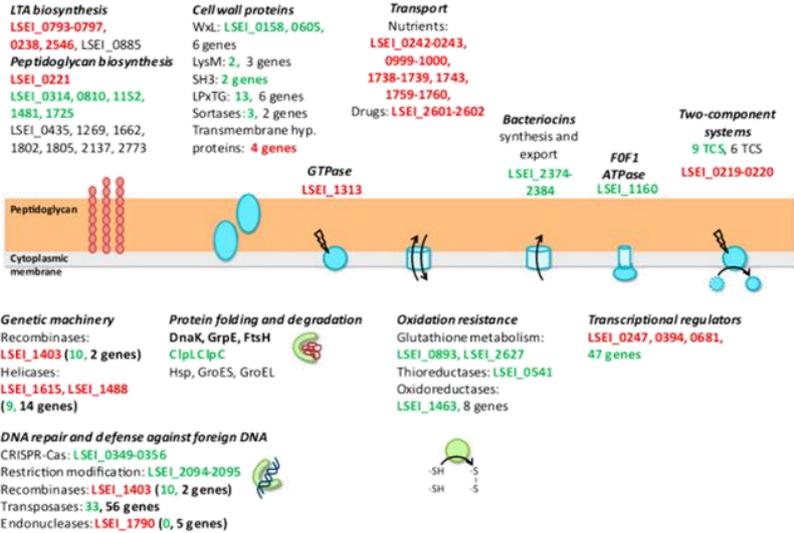


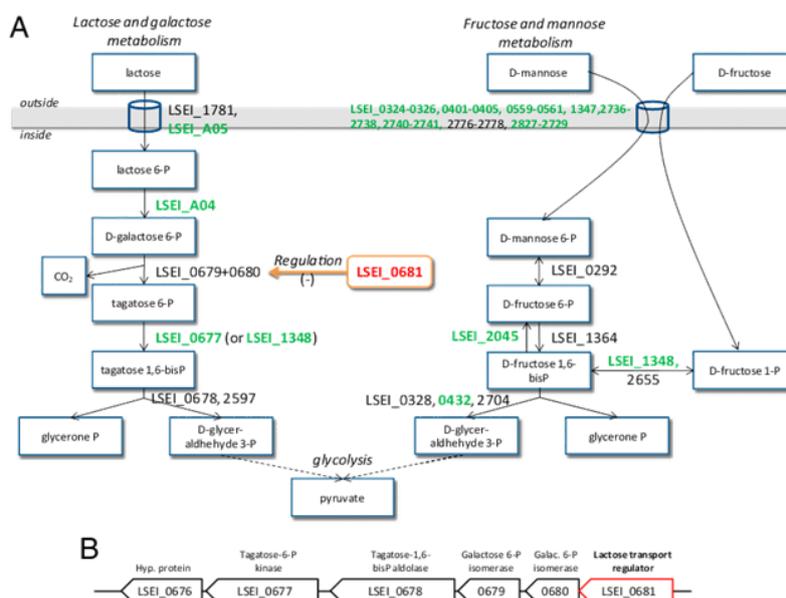
Fig. 3. Main *L. casei* genes for basal metabolism, environmental adaptation, defense mechanisms, cell surface, transport and their role in *L. casei* establishment in the gut. Genes needed for gut persistence are red letters; genes not needed are in green letters; and genes for which no mutant was found in the library (genes not tested) are in black letters.

**Table 1. *L. casei* genes required for intestinal establishment**

Gene tag	Gene name	Operon prediction*
<b>Functions lost in establishment</b>		
Housekeeping		
Cellular machinery		
LSEI_0757	Flavoprotein NrdI	No
LSEI_1274	ATPase for cell division	LSEI_1274-1279 (+)
LSEI_1488	Rad3-related DNA helicase (DinG)	LSEI_1488-1490 (-)
LSEI_1539	Guanosine polyphosphate pyrophosphohydrolase/synthetase	LSEI_1537-1539 (-)
LSEI_1600	RNA-binding protein	LSEI_1600-1601 (-)
LSEI_1615	ATP-dependent DNA helicase RecG	LSEI_1613-1617 (-)
LSEI_1656	tRNA delta (2)-isopentenyl pyrophosphate transferase	No
LSEI_1668	Transcription elongation factor GreA	LSEI_1668-1669 (-)
Adaptation to stress and defense mechanisms		
LSEI_1313	Stress membrane GTPase	No
LSEI_1403	Tyrosine recombinase XerC subunit	LSEI_1402-1403 (+)
LSEI_1790	Predicted Mrr-like endonuclease	No
Metabolism		
Carbohydrate metabolism, transport and energy production		
LSEI_0174	Phosphoketolase	No
LSEI_0681	Lactose transport regulator	LSEI_0676-0681 (-)
LSEI_2549	L-lactate dehydrogenase (LdhL1)	No
Amino acid metabolism		
LSEI_0479	Homoserine transsuccinylase (MetA)	LSEI_0479-0480 (+)
LSEI_0480	Cysteine synthase (CysK)	LSEI_0479-0480 (+)
LSEI_1810	Dipeptidyl aminopeptidase/acylaminoacyl-peptidase	LSEI_1809-1810 (-)
LSEI_2162	Asparagine synthase (AsnB)	LSEI_2161-2162 (-)
Transport		
LSEI_0242	ABC-type Mn/Zn transport system, ATPase component	LSEI_0240-0243 (+)
LSEI_1000	ABC-type Na <sup>+</sup> efflux pump, permease component	LSEI_0999-1000 (+)
LSEI_1738	Peptide ABC transporter permease	LSEI_1738-1739 (-)
LSEI_1743	Cation transport ATPase (CopA)	LSEI_1743-1745 (-)
LSEI_1759	Phosphoenolpyruvate-protein phosphotransferase	LSEI_1759-1760 (-)
LSEI_2601	Multidrug ABC transporter ATPase/permease	LSEI_2599-2602 (+)
Biogenesis of the cell wall		
LSEI_0221	D-alanyl-D-alanine carboxypeptidase	LSEI_0219-0221 (+)
LSEI_0238	PST family polysaccharide transporter	LSEI_0238-0240 (+)
LSEI_0794	D-alanine-activating enzyme (DltA)	LSEI_0793-0797 (+)
LSEI_0796	D-alanyl carrier protein (DltC)	LSEI_0793-0797 (+)
LSEI_0797	D-alanyl transfer protein (DltD)	LSEI_0793-0797 (+)
LSEI_2546	Polysaccharide transporter	LSEI_2545-2546 (-)
Adaptation to environment		
Regulators		
LSEI_0247	Transcriptional regulator	No
LSEI_0394	Transcriptional regulator	LSEI_0394-0396 (+)
LSEI_2025	Fe <sup>2+</sup> /Zn <sup>2+</sup> uptake regulation protein	No
Two-component systems		
LSEI_0219	DNA-binding response regulator	LSEI_0219-0221 (+)
LSEI_0220	Signal transduction histidine kinase	LSEI_0219-0221 (+)
General functions		
Predicted functions		
LSEI_0135	Diadenosine tetraphosphatase-like protein	No
LSEI_0781	Phosphoesterase, DHH family protein	LSEI_0780-0783 (+)
LSEI_1621	GTPase	LSEI_1619-1625 (-)
No predicted functions		
LSEI_0086	Hypothetical protein (transmembrane domain)	LSEI_0084-0086 (+)
LSEI_0156	Hypothetical protein	LSEI_0156-0158 (+)
LSEI_0806	Hypothetical protein (transmembrane domain)	LSEI_0805-0807 (+)
LSEI_1316	Hypothetical protein	LSEI_1314-1319 (+)
LSEI_1461	Hypothetical protein (transmembrane domain)	LSEI_1458-1461 (-)
LSEI_1710	Integral membrane domain	LSEI_1710-1712 (-)
LSEI_2262	Hypothetical protein (102aa)	LSEI_2262-2263 (-)
LSEI_2553	Hypothetical protein	LSEI_2551-2553 (+)
LSEI_A13	Hypothetical protein	No

\*Two consecutive genes were considered to be in the same operon when they were separated by less than 100 bp and no transcription terminator can be predicted. (+) and (-) indicate the orientation of genes, taking the genome sequencing as the (+) orientation.





**Fig. 5.** *L. casei* genes of lactose and mannose metabolism and their role in *L. casei* establishment in the gut. (A) Metabolic pathway context. (B) Genetic environment. Genes needed for gut persistence are in red letters; genes not needed are in green letters; genes for which no mutant was found in the library (genes not tested) are in black letters.

amino acids may be a strong limiting factor in the ileal milieu. The *L. casei* ATCC334 strain can complete only the last step of cysteine synthesis; hence it is auxotrophic for this amino acid. Conversely, mutations in genes for tryptophane, alanine, and threonine synthesis did not affect establishment (Fig. 6 and Table S2). The cysteine synthase activity of the enzyme encoded by LSEI\_0480 was demonstrated experimentally (36). Because LSEI\_0479 and 0480 form an operon (Fig. 6), the phenotype of the LSEI\_0479 mutant could stem from an alteration of LSEI\_0480 transcription alone.

As do all the *Lactobacillus* species, *L. casei* shows numerous auxotrophies for amino acids which would be unfavorable in competitive environments. However, this deficiency is compensated by a large set of peptidases (27 genes) and ABC transporters predicted to transport amino acids or oligopeptides (37). Among the 15 mutated peptidase/protease genes, LSEI\_1810, a dipeptidase predicted to be secreted, is needed for establishment. It may have a high affinity for peptides containing limiting amino acids, such as cysteine and asparagine. Also two ABC transporters predicted to transport peptides, LSEI\_0242-0243 and LSEI\_1738-1739, are needed and could compensate auxotrophies (Fig. 3).

**Environmental adaptation.** Optimal bacterial adaptation to an environment is strongly associated with the expression of a series of sensors/regulators (38, 39). The *L. casei* genome contains 16 complete TCS and 124 transcriptional regulators contributing to its ability to sense and adapt to various environments (37). Among the 10 mutated TCS (a mutant of at least one of the two genes was available in our library) (Table S2), only the LSEI\_0219-0220 TCS, in the same genetic locus as a PBP gene, was needed for gut establishment, reinforcing the importance of the cell wall biogenesis in the ileal context.

Three transcriptional regulators among the 50 whose genes were mutagenized appeared to be required for optimal establishment of *L. casei* (Table S2): LSEI\_0247 (see the discussion of cell wall biogenesis, above), LSEI\_0394 (see the discussion of amino acids metabolism, above), and LSEI\_2025, which belongs to the ferric uptake regulator (Fur) family, a family of metal

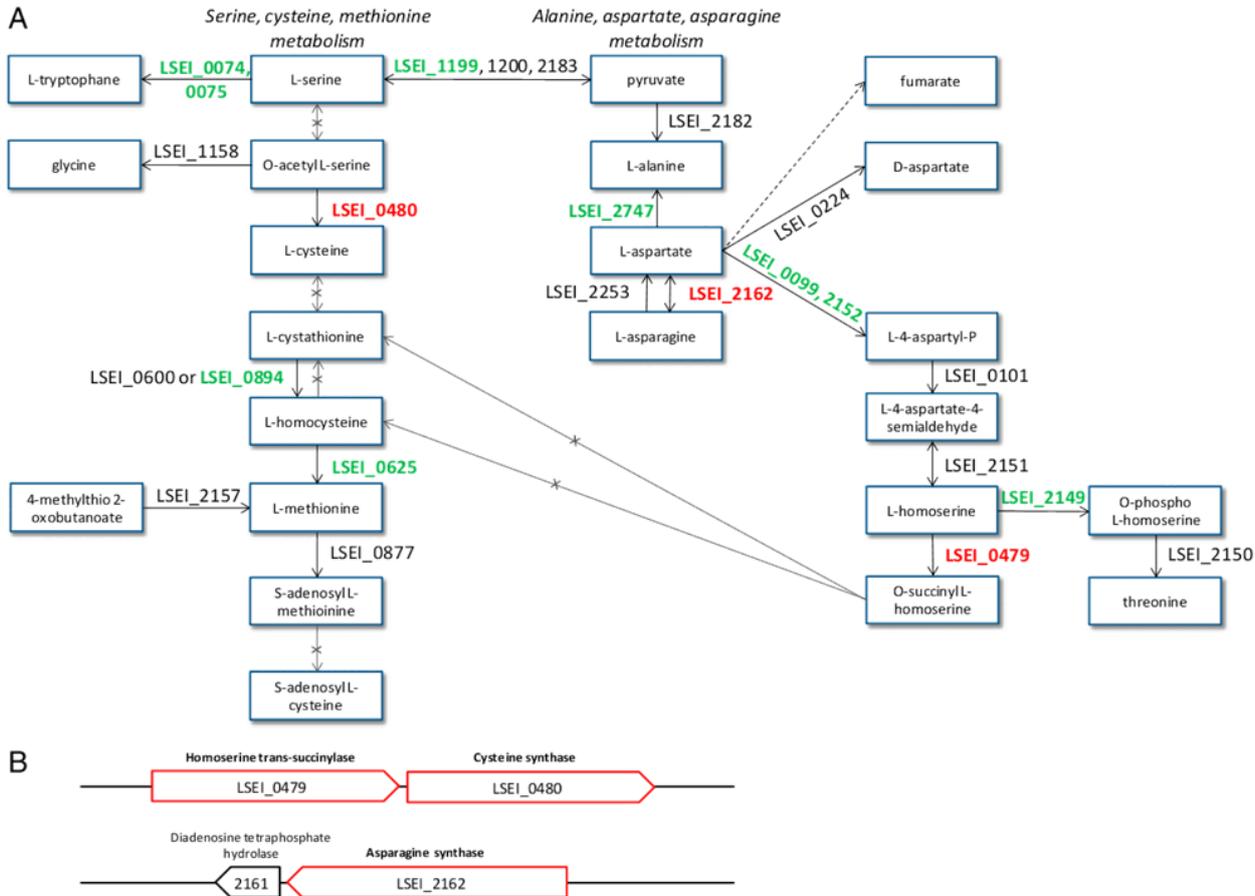
ion-dependent DNA-binding regulators that can sense metal ions (i.e., Fe<sup>2+</sup>, Zn<sup>2+</sup>, Ni<sup>2+</sup>, or Mn<sup>2+</sup>) or hydrogen peroxide.

Otherwise, regularly clustered, interspaced, short palindromic repeats (CRISPR)-Cas genes were reported to regulate gene expression during host-bacteria interaction (40). In this screening, the CRISPR-Cas locus (LSEI\_0349-0356) is dispensable (two mutants were available). Among the six endonucleases of the genome, only one gene, LSEI\_1790, was needed, in contrast to the two genes needed in the type I restriction modification system (LSEI\_2094-2095, a mutant available for each gene) (Table S2). The *L. casei* genome encodes 105 transposases and 13 recombinases, but only one recombinase (LSEI\_1403) was needed for gut establishment. This recombinase could support genetic adaptation to the gut, as shown in *Bacillus fragilis* in which recombinases modulate its surface in the gut (41). Moreover, a single point mutation in the EnvZ/OmpR two-component system affects the expression of more than 100 genes in *Escherichia coli* (42).

## Discussion

Here we report the first, to our knowledge, functional genome-wide study of *L. casei* establishment in the gut. It was made possible by a technological breakthrough that allows transposon-based mutagenesis in this species (20). This approach was combined with differential tagging of the transposon, allowing the assembly of mutant pools that then could be challenged *in vivo*, according to the STM technique that allows the identification of loss-of-function mutants (43, 44). We sequenced the entire set of 9,250 mutants obtained in this study and assembled a library of 1,110 unique mutants in which the tagged transposon had integrated into a discrete gene. Considering that *L. casei* is expected to encode 2,929 genes (45), with a putative set of at least 400 essential genes (46, 47), one can consider that more than half of the genome was mutagenized. Because transposon insertions could have polar effect, one can safely state that nearly all the nonvital *L. casei* functions have been affected.

Our random mutagenesis method focused on identifying the genes necessary for bacterial establishment in the gut regardless



**Fig. 6.** *L. casei* genes of amino acid metabolism and their role in *L. casei* establishment in the gut. (A) Metabolic pathway context. (B) Genetic environment. Genes needed for gut persistence are in red letters; genes not needed are in green letters; genes for which no mutant was found in the library (genes not tested) are in black letters.

the modulation of their in vivo expression, unlike global in vivo transcriptomics [e.g., in vivo expression technology (IVET) or recombination based in vivo expression technology (RIVET)] (11–13). Alternative approaches may call for (i) shotgun cloning of *L. casei* genome fragments into another bacterial species, its success depending upon the basal level of gut establishment of the recipient and the quality of heterospecific gene expression; or (ii) for comparative genomics when the studied strains are phylogenetically close and characterized according to their gut establishment. These approaches identified important genetic loci (14–16) but, unlike random mutagenesis, do not offer an exhaustive genomewide analysis search as does STM. Truly alternative approaches are based on quantification by massive deep sequencing of inoculated pools of thousands of different mutants grown together (17, 48). However, we favored STM followed by mutant quantification using DNA tags. Even though the generation of the mutant library was more labor intensive, it offered significant advantages for our experimental purpose: Mutants were grown individually; pools could be assembled ex-temporaneously; and limiting the pool size allowed a screening model in which the mutants could be exhaustively recovered. Moreover, because of the identification of mutations, mutants now are individually available for dedicated screenings, and the pools also can be assembled differently for other types of studies.

Our first aim was to investigate the functional genomics of the gut-colonization capacity of *L. casei*. We could not carry out

long-term colonization experiments in the mouse intestine because the level of colonization following intragastric administration was insufficient to allow the reliable tracking of each individual mutant composing the injected pool. As an alternative, we validated the rabbit ligated ileal loop model that has been shown to be well adapted to STM screening in *Shigella* (24–26). Because colonization of ligated loops cannot be carried out for more than 16 h, this assay explores the initial survival of mutants to gut luminal conditions and their early steps of colonization, hence the term “establishment.”

**Sensing/Exchange with Environment.** To adapt to ileal conditions, *L. casei* must sense environmental changes to protect itself from toxic compounds while maintaining a functional import of nutrients. Alteration of the cell wall structure may decrease the bacteria’s sensing of its environment and resistance to the environment’s harsh components (49). The consequences may encompass an increase in cellular permeability to environmental compounds and to protons (50), a decrease in resistance to autolysis (32), and a modification of surface properties and hence an enhanced susceptibility to bacteriocins and antimicrobial peptides. Also, some ABC transporters can be involved in resistance to antimicrobial or peptides by excluding these compounds. In fact, ABC transporters, in coordination with a TCS, were shown to be involved in resistance to antimicrobial peptides (51). Also, the mutant in LSEI\_0394, a gene encoding a regulator

of the AcrR family that is located in an operon encoding an ABC transporter (LSEI\_0395-0396) is impaired in its capacity of establishment, whereas the disruption of this transporter had no impact on gut establishment. It is likely that LSEI\_0394 encodes a repressor of this transporter, as is the case for AcrR in *E. coli* on the multidrug efflux pump *acrAB* (52). Therefore, derepression of the transporter becomes a handicap for the bacterium in the gut lumen. The control of exchanges with the ileal environment is fundamental for *L. casei* establishment to capture a maximum of nutrients without unbalancing the bacterial cell content or permitting the entrance of toxic molecules.

**Sugar Metabolism.** Previous transcriptomic studies carried out in *Lactobacillus plantarum* and *L. johnsonii* while colonizing the gut showed an up-regulation of a large set of genes related to carbohydrate transport and metabolism, indicating a global recruitment of sugar-utilization enzymes for energy supply and a change in carbohydrate-utilization pathways to adapt to sugar limitation (15, 53, 54). In *L. plantarum*, energy production from maltose, melibiose, and lactose is activated, as is the import of mannose and cellobiose in the cecum of monocolonized mice (54). Our results are fully consistent with these data. Also, in *B. thetaiotaomicron*, the transcription of numerous genes implicated in polysaccharide degradation is activated in the gut to degrade glycans that the competitive flora (*Bacteroides longum* or *L. casei*) cannot to metabolize (55). Symbionts thus adapt their profile of substrate utilization to the availability of these substrates in the ileal milieu and in response to the presence of other symbionts, suggesting the need to define the composition of the resident microbiota further in future studies.

**Metabolism and Link with Stress/Oxidation.** While initiating colonization, pathogens encounter potentially bactericidal components of the intestinal fluid (i.e., bile, lysozyme, trypsin) and also a strong host response, especially highly reactive oxidative stress molecules, and must respond accordingly to survive. However, it is not clear how much oxidative stress is imposed onto symbionts as they establish in the gut. This screening identified only one general stress protein and no gene encoding factors of the specific response to oxidative stress.

However, cysteine synthase is involved in establishment. Concerning sulfured amino acids, competition for nutrients and during the establishment of a commensal was described for *B. thetaiotaomicron*, particularly the need for vitamin B12 (17), an essential cofactor for methionine biosynthesis for most bacterial species, although not for *L. casei*, according to our genome analysis. Cysteine also constitutes a pool of sulfured molecules implicated in redox regulation of the host intestine and of bacterial cells. The main tandem compounds that permit redox homeostasis are cysteine/cystine, glutathione/glutathione disulfide, and thioredoxin/thioredoxin disulfide (56). Glutathione is synthesized from cysteine in epithelial cells and in some bacteria. *L. casei* is able to use it in complement to thioredoxin to control its redox balance but is unable to synthesize it (57). Because *L. casei* also is auxotrophic for cysteine, it is strongly dependent on the ileal content. Moreover, cysteine is known as the most limiting nonessential amino acid in human cells (58). Thus, *L. casei* must compete with epithelial cells and the endogenous flora for cysteine and glutathione. Because mutants in genes involved in the redox balance, particularly thioreductases, were not affected, we hypothesize that during establishment *L. casei* needs cysteine for nutrition rather than for the maintenance of its redox balance. This notion is consistent with the evidence for nutritional competition, particularly for cysteine, observed between intracellular pathogens and their hosts (59). Although it often involves essential amino acids, nutritional competition also can involve other amino acids, e.g., asparagine, which appears to be decisive for the virulence of *Lactococcus garvieae* (60) and

*Francisella tularensis* (61). Thus, asparagine could be another nutritional requirement for which *L. casei* must compete with its host.

Iron limitation is a major signal in the virulence of mucosal pathogens. In most pathogens Fur proteins act as central regulators for successful colonization and virulence. They control genes involved in iron homeostasis and protection against reactive oxygen species damage (62) in *Shigella* (63), *Salmonella* (64), *Vibrio cholerae* (65), *Pseudomonas* (66), *Listeria* (67), and *Helicobacter pylori* (68). Therefore it is likely that *L. casei* is strongly challenged by iron-limiting conditions in the gut, particularly given the high levels of lactoferrin in intestinal secretions. It also is possible that the protein regulating Fe<sup>2+</sup>/Zn<sup>2+</sup> uptake (LSEI\_2025) allows scavenging of other essential metals and provides *L. casei* protection against reactive oxygen species, although the latter benefit is unlikely because *L. casei* does not appear to activate oxidative stress defenses in the ileal context. Identification of *copA* (LSEI\_1743), an ATPase predicted to be responsible for translocating copper, silver, and cadmium ions across biological membranes, emphasizes the importance of metal import. Accordingly, in *Enterococcus hirae*, *copA* supports bacterial survival in extremely low-copper environments (69). In *L. plantarum*, *copA* expression was highly up-regulated in the conventional mouse gut, and a *copA* mutant showed decreased colonization capacity (70, 71).

In conclusion, we identified 47 mutations affecting establishment that we organized into five functional groups: basic physiological processes (housekeeping), metabolism, cell wall biogenesis, environmental adaptation, and a remaining group of genes of unknown function. In summary, most genes linked to bacterial establishment are conserved among Firmicutes. In consequence, our library provides major information regarding the colonization potential of other Firmicutes. Regarding *L. casei*, we will better characterize the major pathways controlling bacterial establishment. We now are in a position to study the impact of controlled microbiota on the establishment of this model symbiont. Our annotated library also is available to study other phenotypes, particularly the identification of *L. casei* effectors involved in immune and metabolic functions in their colonized host.

## Materials and Methods

**Design of the *L. casei* Random Mutant Library.** For STM, 70 DNA tags, previously used for *Salmonella typhimurium* STM (21) were individually cloned into the EcoRI site of pV1110 to generate 70 differently tagged transposable vectors (Fig. S2 and Tables S3 and S4). The tagged mutant library in *L. casei* was obtained using the P<sub>unc</sub>-T<sub>pas</sub>el<sub>1223</sub> system as recently described (20) and was ordered in pools of 70 mutants. After the transposon insertion sites were identified by individual sequencing, mutants were reassembled to constitute a library of 1,110 gene mutants.

**Screening for Bacterial Establishment.** Each pool of mutants was challenged in rabbit ileal loops as previously described (25, 26) with the following modifications. In each loop, 0.5 mL of bacterial suspension was injected (5 × 10<sup>7</sup> cfu per loop). Challenges were carried out over 16 h. The whole intestinal loop was homogenized, diluted, and spread on agar plates to obtain isolated colonies and to proceed to DNA isolation. Quantitative PCR was used to measure the proportion of each tag corresponding to each mutant in injected and recovered pools. All mutants displaying at least a threefold decrease in quantity between injection and recovery were selected.

Detailed experimental procedures are described in *SI Text*.

**ACKNOWLEDGMENTS.** We thank Ellen Arena for careful reading of and pertinent suggestions regarding the manuscript, Cyril Iaconelli for advice in data processing, Christoph Tang for providing DNA tags, and the Pasteur Institute Micro-organism Collection for the *L. casei* type strain (CIP 107868, ATCC 334). This work and a postdoctoral fellowship (to H.L.-S.) were supported by an Advanced Grant of the European Research Council HOMEOPATH, Grant Agreement ERC-232798 (to P.J.S.). P.J.S. is a Howard Hughes Medical Institute Senior Scholar. H.S. was supported by a doctoral fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche.

1. Savage DC (2001) Microbial biota of the human intestine: A tribute to some pioneering scientists. *Curr Issues Intest Microbiol* 2(1):1–15.
2. O'Hara AM, Shanahan F (2006) The gut flora as a forgotten organ. *EMBO Rep* 7(7):688–693.
3. Brandtzaeg P (2013) Gate-keeper function of the intestinal epithelium. *Benef Microbes* 4(1):67–82.
4. Bocci V (1992) The neglected organ: Bacterial flora has a crucial immunostimulatory role. *Perspect Biol Med* 35(2):251–260.
5. Tremaroli V, Bäckhed F (2012) Functional interactions between the gut microbiota and host metabolism. *Nature* 489(7415):242–249.
6. Eberl G, Boneca IG (2010) Bacteria and MAMP-induced morphogenesis of the immune system. *Curr Opin Immunol* 22(4):448–454.
7. Qin J, et al.; MetaHIT Consortium (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464(7285):59–65.
8. Packey CD, Sartor RB (2009) Commensal bacteria, traditional and opportunistic pathogens, dysbiosis and bacterial killing in inflammatory bowel diseases. *Curr Opin Infect Dis* 22(3):292–301.
9. Arumugam M, et al.; MetaHIT Consortium (2011) Enterotypes of the human gut microbiome. *Nature* 473(7346):174–180.
10. Chow J, Lee SM, Shen Y, Khosravi A, Mazmanian SK (2010) Host-bacterial symbiosis in health and disease. *Adv Immunol* 107:243–274.
11. O'Connell Matherway M, et al. (2011) Functional genome analysis of *Bifidobacterium breve* UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. *Proc Natl Acad Sci USA* 108(27):11217–11222.
12. Bron PA, Grangette C, Mercenier A, de Vos WM, Kleerebezem M (2004) Identification of *Lactobacillus plantarum* genes that are induced in the gastrointestinal tract of mice. *J Bacteriol* 186(17):5721–5729.
13. Walter J, et al. (2003) Identification of *Lactobacillus reuteri* genes specifically induced in the mouse gastrointestinal tract. *Appl Environ Microbiol* 69(4):2044–2051.
14. Kankainen M, et al. (2009) Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human-mucus binding protein. *Proc Natl Acad Sci USA* 106(40):17193–17198.
15. Denou E, et al. (2008) Identification of genes associated with the long-gut-persistence phenotype of the probiotic *Lactobacillus johnsonii* strain NCC533 using a combination of genomics and transcriptome analysis. *J Bacteriol* 190(9):3161–3168.
16. Lee SM, et al. (2013) Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature* 501(7467):426–429.
17. Goodman AL, et al. (2009) Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* 6(3):279–289.
18. Karlsson CL, Molin G, Cilio CM, Ahmés S (2011) The pioneer gut microbiota in human neonates vaginally born at term—a pilot study. *Pediatr Res* 70(3):282–286.
19. Fang F, O'Toole PW (2009) Genetic tools for investigating the biology of commensal lactobacilli. *Front Biosci (Landmark Ed)* 14:3111–3127.
20. Licandro-Seraut H, et al. (2012) Development of an efficient *in vivo* system ( $P_{junc}^{-}$ -TpsaeS<sub>223</sub>) for random transposon mutagenesis of *Lactobacillus casei*. *Appl Environ Microbiol* 78(15):5417–5423.
21. Hensel M, et al. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* 269(5222):400–403.
22. Commichau FM, Pietack N, Stülke J (2013) Essential genes in *Bacillus subtilis*: A re-evaluation after ten years. *Mol Biosyst* 9(6):1068–1075.
23. Caspi R, et al. (2010) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 38(Database issue):D473–D479.
24. Sansonetti PJ, et al. (1983) Alterations in the pathogenicity of *Escherichia coli* K-12 after transfer of plasmid and chromosomal genes from *Shigella flexneri*. *Infect Immun* 39(3):1392–1402.
25. Marteyn B, et al. (2010) Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature* 465(7296):355–358.
26. West NP, et al. (2005) Optimization of virulence functions through glycosylation of *Shigella* LPS. *Science* 307(5713):1313–1317.
27. Oozeer R, Mater DD, Goupil-Feuillerat N, Corthier G (2004) Initiation of protein synthesis by a labeled derivative of the *Lactobacillus casei* DN-114 001 strain during transit from the stomach to the cecum in mice harboring human microbiota. *Appl Environ Microbiol* 70(12):6992–6997.
28. McRobbie AM, et al. (2012) *Staphylococcus aureus* DinG, a helicase that has evolved into a nuclease. *Biochem J* 442(1):77–84.
29. Alcántara C, Zúñiga M (2012) Proteomic and transcriptomic analysis of the response to bile stress of *Lactobacillus casei* BL23. *Microbiology* 158(Pt 5):1206–1218.
30. Kleerebezem M, et al. (2010) The extracellular biology of the lactobacilli. *FEMS Microbiol Rev* 34(2):199–230.
31. Kawai Y, et al. (2011) A widespread family of bacterial cell wall assembly proteins. *EMBO J* 30(24):4931–4941.
32. Walter J, et al. (2007) D-alanyl ester depletion of teichoic acids in *Lactobacillus reuteri* 100-23 results in impaired colonization of the mouse gastrointestinal tract. *Environ Microbiol* 9(7):1750–1760.
33. Rico J, Yebra MJ, Pérez-Martínez G, Deutscher J, Monedero V (2008) Analysis of *ldh* genes in *Lactobacillus casei* BL23: Role on lactic acid production. *J Ind Microbiol Biotechnol* 35(6):579–586.
34. Viana R, Yebra MJ, Galán JL, Monedero V, Pérez-Martínez G (2005) Pleiotropic effects of lactate dehydrogenase inactivation in *Lactobacillus casei*. *Res Microbiol* 156(5-6):641–649.
35. Viana R, et al. (2000) Enzyme I and Hpr from *Lactobacillus casei*: Their role in sugar transport, carbon catabolite repression and inducer exclusion. *Mol Microbiol* 36(3):570–584.
36. Bogicevic B, Berthoud H, Portmann R, Meile L, Irmeler S (2012) CysK from *Lactobacillus casei* encodes a protein with O-acetylserine sulfhydrylase and cysteine desulfurization activity. *Appl Microbiol Biotechnol* 94(5):1209–1220.
37. Cai H, Thompson R, Budinich MF, Broadbent JR, Steele JL (2009) Genome sequence and comparative genome analysis of *Lactobacillus casei*: Insights into their niche-associated evolution. *Genome Biol Evol* 1:239–257.
38. Sturme MH, Francke C, Siezen RJ, de Vos WM, Kleerebezem M (2007) Making sense of quorum sensing in lactobacilli: A special focus on *Lactobacillus plantarum* WCFS1. *Microbiology* 153(Pt 12):3939–3947.
39. Krell T, et al. (2010) Bacterial sensor kinases: Diversity in the recognition of environmental signals. *Annu Rev Microbiol* 64:539–559.
40. Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL, Weiss DS (2013) A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497(7448):254–257.
41. Coyne MJ, Weinacht KG, Krinos CM, Comstock LE (2003) Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. *Proc Natl Acad Sci USA* 100(18):10446–10451.
42. Giraud A, et al. (2008) Dissecting the genetic components of adaptation of *Escherichia coli* to the mouse gut. *PLoS Genet* 4(1):e2.
43. Saenz HL, Dehio C (2005) Signature-tagged mutagenesis: Technical advances in a negative selection method for virulence gene identification. *Curr Opin Microbiol* 8(5):612–619.
44. West NP, Sansonetti PJ, Frankel G, Tang CM (2003) Finding your niche: What has been learnt from STM studies on GI colonization. *Trends Microbiol* 11(7):338–344.
45. Makarova K, et al. (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci USA* 103(42):15611–15616.
46. Acevedo-Rocha CG, Fang G, Schmidt M, Ussery DW, Danchin A (2013) From essential to persistent genes: A functional approach to constructing synthetic life. *Trends Genet* 29(5):273–279.
47. Juhas M, Eberl L, Glass JI (2011) Essence of life: Essential genes of minimal genomes. *Trends Cell Biol* 21(10):562–568.
48. Langridge GC, et al. (2009) Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res* 19(12):2308–2316.
49. Jordan S, Hutchings MI, Mascher T (2008) Cell envelope stress response in Gram-positive bacteria. *FEMS Microbiol Rev* 32(1):107–146.
50. Boyd DA, et al. (2000) Defects in D-alanyl-lipoteichoic acid synthesis in *Streptococcus mutans* results in acid sensitivity. *J Bacteriol* 182(21):6055–6065.
51. Revilla-Guarinos A, et al. (2013) Characterization of a regulatory network of peptide antibiotic detoxification modules in *Lactobacillus casei* BL23. *Appl Environ Microbiol* 79(10):3160–3170.
52. Routh MD, Su CC, Zhang Q, Yu EW (2009) Structures of AcrR and CmeR: Insight into the mechanisms of transcriptional repression and multi-drug recognition in the TetR family of regulators. *Biochim Biophys Acta* 1794(5):844–851.
53. Denou E, et al. (2007) Gene expression of commensal *Lactobacillus johnsonii* strain NCC533 during *in vitro* growth and in the murine gut. *J Bacteriol* 189(22):8109–8119.
54. Marco ML, et al. (2009) Lifestyle of *Lactobacillus plantarum* in the mouse caecum. *Environ Microbiol* 11(10):2747–2757.
55. Sonnenburg JL, Chen CT, Gordon JI (2006) Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host. *PLoS Biol* 4(12):e143.
56. Circa ML, Aw TY (2012) Intestinal redox biology and oxidative stress. *Semin Cell Dev Biol* 23(7):729–737.
57. Serata M, Iino T, Yasuda E, Sako T (2012) Roles of thioredoxin and thioredoxin reductase in the resistance to oxidative stress in *Lactobacillus casei*. *Microbiology* 158(Pt 4):953–962.
58. Young VR (1994) Adult amino acid requirements: The case for a major revision in current recommendations. *J Nutr* 124(8, Suppl):1517S–1523S.
59. Abu Kwaik Y, Bumam D (2013) Microbial quest for food in vivo: 'Nutritional virulence' as an emerging paradigm. *Cell Microbiol* 15(6):882–890.
60. Menéndez A, Fernández L, Reimundo P, Guijarro JA (2007) Genes required for *Lactococcus garvieae* survival in a fish host. *Microbiology* 153(Pt 10):3286–3294.
61. Gesbert G, et al. (2014) Asparagine assimilation is critical for intracellular replication and dissemination of Francisella. *Cell Microbiol* 16(3):434–449.
62. Troxell B, Hassan HM (2013) Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Front Cell Infect Microbiol* 3:59.
63. Africa LA, Murphy ER, Egan NR, Wigley AF, Wing HJ (2011) The iron-responsive Fur/RyhB regulatory cascade modulates the *Shigella* outer membrane protease IcsP. *Infect Immun* 79(11):4543–4549.
64. Leclerc JM, Dozois CM, Daigle F (2013) Role of the *Salmonella enterica* serovar Typhi Fur regulator and small RNAs RfrA and RfrB in iron homeostasis and interaction with host cells. *Microbiology* 159(Pt 3):591–602.
65. Mey AR, Wyckoff EE, Kanukurthy V, Fisher CR, Payne SM (2005) Iron and fur regulation in *Vibrio cholerae* and the role of fur in virulence. *Infect Immun* 73(12):8167–8178.
66. Cornelis P, Matthijs S, Van Oeffelen L (2009) Iron uptake regulation in *Pseudomonas aeruginosa*. *Biomol* 22(1):15–22.
67. McLaughlin HP, et al. (2012) A putative P-type ATPase required for virulence and resistance to haem toxicity in *Listeria monocytogenes*. *PLoS ONE* 7(2):e30928.
68. Pich OQ, Merrell DS (2013) The ferric uptake regulator of *Helicobacter pylori*: A critical player in the battle for iron and colonization of the stomach. *Future Microbiol* 8(6):725–738.
69. Solioz M, Stoyanov JV (2003) Copper homeostasis in *Enterococcus hirae*. *FEMS Microbiol Rev* 27(2-3):183–195.
70. Marco ML, Bongers RS, de Vos WM, Kleerebezem M (2007) Spatial and temporal expression of *Lactobacillus plantarum* genes in the gastrointestinal tracts of mice. *Appl Environ Microbiol* 73(1):124–132.
71. Bron PA, Meijer M, Bongers RS, de Vos WM, Kleerebezem M (2007) Dynamics of competitive population abundance of *Lactobacillus plantarum* *ivi* gene mutants in faecal samples after passage through the gastrointestinal tract of mice. *J Appl Microbiol* 103(5):1424–1434.
72. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40(Database issue):D109–D114.

# Supporting Information

Licandro-Seraut et al. 10.1073/pnas.1411883111

## SI Text

**Determination of the Relevant Threshold to Screen the Decrease of Mutants After in Vivo Challenge.** To determine the relevant threshold of variation between the injected and recovered pools, we challenged a pool of 60 mutants in two rabbits and quantified these 60 tagged mutants in the injected pool and in the two recovered pools. The mean variation (the ratio of the quantity of injected mutants to the quantity of recovered mutants) was  $1.13 \pm 0.37$  for the first rabbit and  $1.01 \pm 0.33$  for the second. The maximal decrease was less than 2.5-fold; Fig. S2 presents the distribution of these variations. Therefore we decided to use a threefold decrease in the recovered mutants relative to injected mutants as the threshold in screening the mutant gene library.

**Screening of 3,000 Unidentified, Randomly Selected Tagged Mutants in the Rabbit Ligated Ileal Loop Model.** We carried out our initial mutant screening on random, unidentified mutants, considering that all mutants could be informative, whether the transposon target was a gene or an intergenic region. Indeed, intergenic regions can contain regulatory elements that also may affect bacterial adaptation. Moreover, studying mutants in intergenic regions is a way to target essential genes, because mutation in their promoter or transcriptional terminator can modify their transcription without completely knocking out function.

Forty-three pools of 70 tagged mutants were challenged in rabbit ligated ileal loops (one pool per loop,  $5.10^7$  cfu per loop) for a duration of 16 h. Because we anticipated that the mechanisms implicated in gut establishment are complex and multi-genic, we applied a low-stringency threshold. The proportion of each mutant in the recovered pool was determined by quantitative PCR (qPCR) using the 70 different DNA tags. When the proportion of a mutant in the recovered pool was at least fivefold less or fivefold more than 1/70, we used qPCR to determine the proportion of this mutant in the injected pool. Each identified mutant was challenged again in a pool of 70 mutants, and mutants displaying at least a threefold decrease or increase in quantity between the input and output pools in the two independent screenings were selected. Several genes were identified in this screening, and all of them also were identified in the screening of gene mutants. Only two mutants in an intergenic region showed alterations in gut establishment. The first mutant was a transposon insertion in the LSEI\_1600 gene transcription terminator. Because the LSEI\_1600 gene already had been identified as a gut establishment gene (*SI Materials and Methods* and Table 1), the gut-sensitive phenotype of this intergenic mutant probably results from a decrease of LSEI\_1600 mRNA stability. For the second mutant, the transposon was integrated in a 660-bp intergenic region between genes LSEI\_0318 and LSEI\_0319. This sequence does not match noncoding RNA in the Rfam database (1). This initial screening revealed that the intergenic region contributed very little to gut establishment; therefore we restricted our search to mutants in which transposons were inserted into genes.

## SI Materials and Methods

**Bacterial Strains and Culture Conditions.** *Lactobacillus casei* CIP 107868 (ATCC 334) was grown statically in MRS medium (Difco) at 37 °C. *Escherichia coli* TG1 was grown aerobically in LB medium with shaking at 37 °C. The antibiotics used were 50 µg/mL ampicillin and 50 µg/mL erythromycin for *E. coli* and 5 µg/mL erythromycin and 5 µg/mL chloramphenicol for *L. casei*.

**Genetic Manipulation.** DNA manipulation, purification, ligation, restriction analysis, and gel electrophoresis were carried out as described by Sambrook et al. (2). PCR and restriction products were purified using a gel and PCR clean-up kit (Macherey Nagel). Plasmids were prepared from *E. coli* using the Nucleospin plasmid kit (Macherey Nagel). *L. casei* and *E. coli* were transformed by electroporation as described by Aukrust et al. (3) and Dower et al. (4), respectively, using a Gene Pulser and a Pulse Controller apparatus (Bio-Rad).

To proceed to signature-tagged mutagenesis (STM), 70 DNA tags previously used for *Salmonella typhimurium* STM (5), composed of invariable arms with EcoRI and HindIII restriction sites surrounding a variable region about 40-bp long, were cloned individually into the EcoRI site of pV1110 to generate 70 differently tagged transposable vectors (Fig. S1). Tag sequence was determined by sequencing (GATC Biotech) using primers PVI110TAG1 (5'-GATCAATCTTGAAGACGAAAG-3') and PVI110TAG2 (5'-CTGCAAGGCGATTAAGTTGG-3') that anneal 100 nt upstream and downstream of the tag insertion site, respectively.

The tagged mutant library in *L. casei* was obtained using the  $P_{\text{junc}}\text{-TpaseIS}_{1223}$  system as recently described (6). Briefly, *L. casei* first was transformed with an unstable plasmid (pV1129) carrying the IS<sub>1223</sub> transposase gene and then was transformed independently with the 70 tagged transposable vectors generated in this study (the integrative pV1110 plasmid carrying the transposon and an erythromycin-resistance cassette). Mutants obtained for each tagged transposable vector were ordered in plates.

**Sequencing the Transposon Insertion Sites and Assembling the Library of ORF Mutants.** Genomic DNA was extracted individually from each mutant using the NucleoSpin 96 Tissue kit (Macherey Nagel) with some modifications adapted for *Lactobacillus*. Each DNA extract using the primer IRL (5'-GTTGG-GAGCTCTCCCATATG-3') was sequenced to identify transposon insertion sites, and the obtained sequences were BLASTed against the *L. casei* genome. Thereafter, mutants were sorted by the localization of the transposon. The mutants in which the transposon was integrated into a gene for which a mutant had not already been found were reassembled in pools. Because all mutants in the pools were tagged differently, the number of mutants in each pool ranged from 25 to 70. When there were fewer than 40 mutants in a pool, intergenic region mutants were added to attain the 40 mutants needed to maintain sufficient competition between mutants.

**Inoculation of Rabbit Ligated Ileal Loops.** Mutants were grown individually in 96-well microplates, assembled in pools, washed with PBS, and resuspended in PBS to a final OD<sub>600</sub> of 0.1. They were challenged in rabbit ligated ileal loops as previously described (7, 8) with the following modifications. Between 10 and 12 ileal loops were made in each rabbit. In each loop, 0.5 mL of bacterial suspension was injected (corresponding to a total bacterial quantity of  $5 \times 10^7$  cfu per loop). Challenges were carried out over 16 h. Rabbit were killed, and each loop was removed integrally by severing the spacers just outside the ligatures and was transferred into cold PBS. The whole intestinal loop containing the bacteria then was homogenized by bead-beating using 2.8-mm ceramic beads in a Precellys 24 apparatus. The resulting homogenate (about  $5 \geq 10^7$  living *L. casei* bacteria per loop) was diluted and spread on agar plates to obtain isolated colonies. Each pool was screened in two different rabbits, and the position of injected pools was randomized. This in vivo model protocol was approved

by the Comité Régional d’Ethique pour l’Experimentation Animale in Paris 1 (protocol no. 20070004).

**Mutant Quantification After in Vivo Challenge.** Genomic DNA was isolated from injected and recovered pools. For each injected pool, 5 mL of cell culture was used. For each recovered pool, 5,000–10,000 colonies were recovered by washing plates with TE buffer (5 mM Tris, 2 mM EDTA). Then the same DNA isolation procedure was applied. Bacterial suspensions were washed with TE buffer and centrifuged. Bacterial pellets were incubated for 1 h at 65 °C in 500 µL lysis buffer (TE buffer, 1% SDS, 0.4 mg/mL proteinase K). The obtained cell suspensions were mixed with 0.7 mL phenol:chloroform:isoamyl alcohol at a ratio of 25:24:1 and with 0.1-µm glass beads and were disrupted by bead-beating using one treatment of 40 s at a speed of 6.5 in a FastPrep reciprocating device (Qbiogene). DNAs were isolated from the aqueous phase by two phenol-chloroform extractions, followed by DNA precipitation in ethanol and washing in 70% ethanol. DNA pellets were dissolved in 200 µL TE buffer.

qPCR was used to measure the proportion of each mutant in the pools. Based on the tag sequences (Table S3), primers used for qPCR were designed using the Primer3 software (Table S4). The first primer corresponded to an invariable pV110 region and was common to the 70 primer couples. The second primers were designed on DNA tags to generate a 105- to 125-bp amplicon when associated with the first primer (Fig. S2). qPCR amplifications were carried out on a Bio-Rad iCycler. Each 20-µL reaction contained diluted isolated DNA, 0.25 µM of each

primer, and 10 µL iQ SYBR Green Supermix (Bio-Rad). The QPVITAGR primer was associated with a tag-specific primer (Table S4) to create a specific primer couple for the enumeration of each tag. The 70 tagged transposable vectors were pooled at equimolar concentration, and serial dilutions of this mix were used as the matrix for qPCR with each primer couple to check the correlation with the initial DNA concentration. Four sets of six randomly selected tagged transposable vectors were diluted in different known concentrations and subjected to real-time quantification. For these 24 tagged transposable vectors, relative quantification by real-time PCR correlated with the concentration, confirming that amplifications were tag-specific. Primer couples CTRL3F-CTRL3R and CTRL4F-CTRL4R, designed on the tagged-pV110 invariable region, were used for each mutant pool for global enumeration of all tags. Tags were quantified using comparative quantification (Cq) methods. The mean Cq values obtained for the CTRL3F-CTRL3R and CTRL4F-CTRL4R amplicons were used to determine the ratio of each tag to the complete set of tags in the injected and recovered pools. The decreased rate for each mutant corresponds to the ratio of the corresponding tag in the injected pool divided by the ratio in the recovered pool.

**Bioinformatics.** Gene annotations were found using the Integrated Microbial Genomes site (9) based on the analysis by Makarova et al. (10). Gene function predictions were made using the Kyoto Encyclopedia of Genes and Genomes (11) and Biocyc (12) databases.

- Burge SW, et al. (2013) Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res* 41: D226–D232.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY) 2nd Ed.
- Aukrust TW, Brurberg MB, Nes IF (1995) Transformation of *Lactobacillus* by electroporation. *Methods Mol Biol* 47:201–208.
- Dower WJ, Miller JF, Ragsdale CW (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* 16(13):6127–6145.
- Hensel M, et al. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* 269(5222):400–403.
- Licandro-Seraut H, et al. (2012) Development of an efficient *in vivo* system ( $P_{junc}$ -T<sub>psalS1,223</sub>) for random transposon mutagenesis of *Lactobacillus casei*. *Appl Environ Microbiol* 78(15):5417–5423.
- Marteyn B, et al. (2010) Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature* 465(7296):355–358.
- West NP, et al. (2005) Optimization of virulence functions through glycosylation of *Shigella* LPS. *Science* 307(5713):1313–1317.
- Markowitz VM, et al. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Res* 34(Database issue):D344–D348.
- Makarova K, et al. (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci USA* 103(42):15611–15616.
- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40(Database issue): D109–D114.
- Caspi R, et al. (2010) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 38(Database issue):D473–D479.

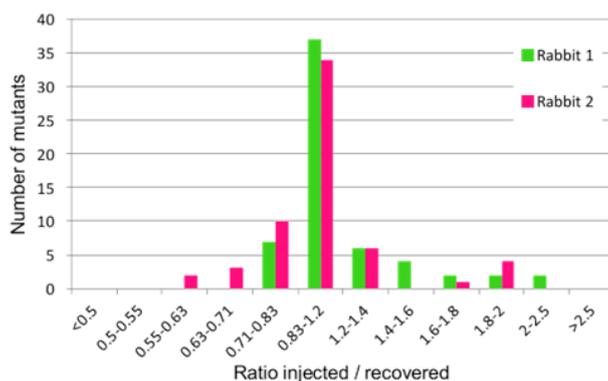


Fig. S1. Distribution of variation in the ratio of the relative quantity of each mutant in the injected pool vs. the recovered pool. Analysis was made from one pool of 60 mutants injected in two rabbits.

Table S1. Numeration of CFU per DO unit for cultures in stationary phase of growth.

Mutant	CFU / uDO	Mutant	CFU / uDO
LSEI_0086	3,88E+08	<b>LSEI_1468</b>	<b>2,57E+07</b>
LSEI_0135	2,82E+08	LSEI_1488	2,45E+08
LSEI_0156	2,23E+08	<b>LSEI_1565</b>	<b>3,35E+07</b>
LSEI_0186	2,00E+08	LSEI_1566	4,52E+08
LSEI_0174	2,71E+08	LSEI_1580	1,92E+08
LSEI_0219	1,88E+08	LSEI_1600	2,82E+08
LSEI_0220	2,04E+08	LSEI_1615	2,42E+08
LSEI_0221	3,04E+08	LSEI_1621	3,14E+08
LSEI_0238	1,94E+08	LSEI_1656	2,63E+08
LSEI_0242	5,49E+08	LSEI_1659	3,23E+08
LSEI_0247	2,14E+08	LSEI_1668	2,25E+08
<b>LSEI_0281</b>	<b>1,96E+07</b>	LSEI_1677	3,66E+08
LSEI_0394	1,51E+08	LSEI_1684	2,55E+08
LSEI_0479	5,16E+08	LSEI_1710	1,98E+08
LSEI_0480	5,00E+08	LSEI_1738	1,88E+08
LSEI_0517	1,16E+08	LSEI_1740	2,23E+08
LSEI_0625	3,26E+08	LSEI_1743	2,57E+08
LSEI_0681	3,55E+08	LSEI_1759	2,47E+08
LSEI_0728	2,33E+08	LSEI_1790	2,69E+08
LSEI_0756	1,77E+08	LSEI_1810	2,86E+08
LSEI_0757	1,94E+08	LSEI_2005	2,51E+08
LSEI_0781	2,86E+08	LSEI_2009	3,93E+08
LSEI_0794	9,85E+07	LSEI_2025	2,59E+08
LSEI_0796	1,13E+08	LSEI_2162	3,91E+08
LSEI_0797	9,87E+07	LSEI_2236	1,67E+08
LSEI_0806	1,63E+08	LSEI_2269	2,65E+08
LSEI_1000	1,26E+08	LSEI_2349	1,68E+08
LSEI_1156	2,99E+08	LSEI_2540	3,36E+08
LSEI_1274	3,83E+08	LSEI_2546	2,15E+08
<b>LSEI_1278</b>	<b>1,35E+07</b>	LSEI_2553	3,37E+08
LSEI_1293	1,85E+08	LSEI_2829	3,21E+08
LSEI_1313	6,54E+08	LSEI_A13	3,08E+08
LSEI_1348	2,02E+08	LSEI_A16	1,95E+08
LSEI_1403	1,70E+08	LSEI_r0259	3,29E+08
LSEI_1461	4,10E+08	WT	2,70E+08

Mean value for 3 independent experiments. Mean variation is 10%. Mutants for which CFU/OD unit ratio is less than 5-fold the WT ratio are in bold.

Table S2. Genes involved in the discussed functional groups that are not essential for gut persistence.

Regulatory functions and adaptation.		COG number
Locus tag	Name	
<i>Two-component systems</i>		
LSEI_0460	Response regulator consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	COG0745
LSEI_0934	Response regulator consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	COG0745
LSEI_0935	Signal transduction histidine kinase	COG5002
LSEI_1042	Signal transduction histidine kinase	COG0642
LSEI_1223	Signal transduction histidine kinase	COG4585
LSEI_1419	Signal transduction histidine kinase	COG0642
LSEI_1678	Signal transduction histidine kinase	COG0642
LSEI_1740	Response regulator consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	COG0745
LSEI_2680	Signal transduction histidine kinase	COG0642
LSEI_2868	Signal transduction histidine kinase regulating citrate/malate metabolism	COG3290
<i>Transcriptional regulators</i>		
LSEI_0023	Transcriptional regulator	COG1309
LSEI_0052	Transcriptional regulator	COG1802
LSEI_0061	Transcriptional regulator	COG1309
LSEI_0081	Transcriptional regulator	COG1309
LSEI_0110	Transcriptional regulator	COG1309
LSEI_0121	Transcriptional regulator	COG1737
LSEI_0132	Predicted transcriptional regulator	COG1476
LSEI_0222	Transcription elongation factor	COG0782
LSEI_0247	Transcriptional regulator	COG1316
LSEI_0250	Transcriptional regulator of sugar metabolism	COG1349
LSEI_0297	Putative transcription activator	COG0819
LSEI_0330	Predicted transcriptional regulator	COG1378
LSEI_0337	Transcriptional regulator	COG1309
LSEI_0368	Transcriptional regulator	COG1609
LSEI_0371	Transcriptional regulator	COG1737
LSEI_0428	Predicted transcriptional regulator	COG0640
LSEI_0453	Transcriptional antiterminator	COG3711
LSEI_0473	Transcriptional regulator	COG0583
LSEI_0558	Transcriptional antiterminator	COG3933
LSEI_0681	Transcriptional regulator of sugar metabolism	COG1349
LSEI_0739	Transcriptional regulator	COG0583
LSEI_0802	Predicted transcriptional regulator with C-terminal CBS domains	COG3620
LSEI_1147	Transcriptional regulator	COG1309

LSEI_1240	Transcriptional accessory protein	COG2183
LSEI_1243	Predicted transcriptional regulator	COG1395
LSEI_1413	Transcriptional regulator	COG0583
LSEI_1471	Transcriptional regulator	COG0583
LSEI_1659	Transcriptional regulator/sugar kinase	COG1940
LSEI_1668	Transcription elongation factor	COG0782
LSEI_2033	Predicted transcriptional regulator	COG0640
LSEI_2050	Transcriptional regulator	COG1316
LSEI_2127	Transcriptional regulator	COG1846
LSEI_2189	Transcriptional regulator	COG1309
LSEI_2244	Predicted transcriptional regulator	COG1725
LSEI_2269	Transcriptional regulator	COG1609
LSEI_2370	Predicted transcriptional regulator	COG1396
LSEI_2371	Predicted transcriptional regulator	COG1695
LSEI_2464	Transcriptional regulator	COG1309
LSEI_2468	Predicted transcriptional regulator	COG0789
LSEI_2670	Transcriptional regulator	COG1414
LSEI_2719	Transcriptional antiterminator	COG3711
LSEI_2724	Glucitol operon activator	COG4578
LSEI_2725	Transcriptional antiterminator	COG3711
LSEI_2728	Transcriptional antiterminator	COG3711
LSEI_2734	Transcriptional regulator of sugar metabolism	COG1349
LSEI_2742	Transcriptional antiterminator	COG3711
LSEI_2759	Predicted transcriptional regulator	COG1476
LSEI_A02	Predicted transcriptional regulator	COG3682
LSEI_A06	Transcriptional antiterminator	COG3711

**Stress response**

Locus tag	Name	COG number
<i>stress related enzymes</i>		
LSEI_0206	Pyroglutamate-carboxylate peptidase (N-terminal pyroglutamyl peptidase)	COG2039
LSEI_0255	ADP-ribosylglycohydrolase	COG1397
LSEI_0541	Thioredoxin domain-containing protein	COG3118
LSEI_0638	Predicted redox protein, regulator of disulfide bond formation	COG1764
LSEI_0742	Peroxioredoxin	COG2077
LSEI_0827	Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family	COG0652
LSEI_0893	Glutathione peroxidase	COG0386
LSEI_1202	Peptide methionine sulfoxide reductase	COG0225
LSEI_1393	Peptide methionine sulfoxide reductase	COG0225
LSEI_1412	Pyruvate-formate lyase-activating enzyme	COG1180
LSEI_1463	Predicted oxidoreductase related to nitroreductase	COG3560
LSEI_1472	Predicted metalloendopeptidase	COG3590
LSEI_1525	Conserved domain frequently associated with peptide methionine sulfoxide reductase	COG0229
LSEI_1762	ATPases with chaperone activity, ATP-binding subunit	COG0542
LSEI_2048	ATPases with chaperone activity, ATP-binding subunit	COG0542
LSEI_2270	Subtilisin-like serine proteases	COG1404
LSEI_2517	ATPases with chaperone activity, ATP-binding subunit	COG0542
LSEI_2627	Glutathione-disulfide reductase	COG1249
<i>ABC-type transporters and efflux pump</i>		
LSEI_0057	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_0091	ABC-type multidrug transport system, ATPase component	COG1131
LSEI_0106	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_0129	ABC-type multidrug transport system, ATPase component	COG1131
LSEI_0178	ABC-type multidrug transport system, ATPase component	COG1131
LSEI_0183	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_0395	ABC-type antimicrobial peptide transport system, permease component	COG0577
LSEI_0620	ABC-type multidrug transport system, ATPase component	COG1131
LSEI_0626	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_0627	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_1014	ABC-type multidrug transport system, permease component	COG0842
LSEI_1148	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_1149	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_1180	ABC-type transport system involved in Fe-S cluster assembly, ATPase component	COG0396
LSEI_1418	ABC-type antimicrobial peptide transport system, ATPase component	COG1136
LSEI_1814	ABC-type antimicrobial peptide transport system, ATPase component	COG1136
LSEI_1873	ABC-type multidrug transport system, ATPase component	COG1131
LSEI_1884	ABC-type antimicrobial peptide transport system, ATPase component	COG1136
LSEI_1994	ABC-type antimicrobial peptide transport system, ATPase component	COG1136
LSEI_2194	ABC-type multidrug transport system, ATPase and permease components	COG1132
LSEI_2322	ABC-type multidrug transport system, ATPase component	COG1131

LSEI_2332	ABC-type multidrug transport system, ATPase component	COG1131		
LSEI_2384	ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain	COG2274		
LSEI_2466	ABC-type antimicrobial peptide transport system, ATPase component	COG1136		
LSEI_2602	ABC-type multidrug transport system, ATPase and permease components	COG1132		
LSEI_2621	Na <sup>+</sup> -driven multidrug efflux pump	COG0534		
LSEI_2626	ABC-type antimicrobial peptide transport system, ATPase component	COG1136		
LSEI_2894	ABC-type antimicrobial peptide transport system, ATPase component	COG1136		
LSEI_2895	ABC-type antimicrobial peptide transport system, permease component	COG0577		
	<b>DNA recombination and repair</b>			
LSEI_0018	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0037	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0050	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0092	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0237	Transposase and inactivated derivatives	COG3464		
LSEI_0272	Transposase and inactivated derivatives	COG2963		
LSEI_0311	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0344	Site-specific recombinases, DNA invertase Pin homologs	COG1961		
LSEI_0345	Transposase and inactivated derivatives	COG3293		
LSEI_0347	Transposase and inactivated derivatives	COG3316		
LSEI_0356	DNA polymerase III, alpha subunit (gram-positive type)	COG2176		
LSEI_0361	Transposase and inactivated derivatives	COG3464		
LSEI_0413	Transposase and inactivated derivatives	COG2801		
LSEI_0446	Transposase and inactivated derivatives	COG3464		
LSEI_0454	Site-specific recombinases, DNA invertase Pin homologs	COG1961		
LSEI_0486	Site-specific recombinase XerD	COG4974		
LSEI_0507	Putative primosome component and related proteins	COG3935		
LSEI_0508	DNA replication protein	COG1484		
LSEI_0516	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0570	Adenine specific DNA methylase Mod	COG2189		
LSEI_0573	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0579	Site-specific recombinases, DNA invertase Pin homologs	COG1961		
LSEI_0586	Adenine specific DNA methylase Mod	COG2189		
LSEI_0587	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0590	Transposase and inactivated derivatives	COG2801		
LSEI_0684	ATPase involved in DNA repair	COG0419		
LSEI_0689	A/G-specific DNA glycosylase	COG1194		
LSEI_0773	Transposase and inactivated derivatives	COG2801		
LSEI_0774	Transposase and inactivated derivatives	COG2963		
LSEI_0791	Mismatch repair ATPase (MutS family)	COG1193		
LSEI_0857	Transposase and inactivated derivatives	COG3436		
LSEI_0907	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_0923	RecA/RadA recombinase	COG0468		
LSEI_0957	Excinuclease ATPase subunit	COG0178		
LSEI_0995	Uracil DNA glycosylase	COG0692		
LSEI_1001	DNA polymerase III, epsilon subunit and related 3'-5' exonucleases	COG0847		
LSEI_1049	Nucleotidyltransferase/DNA polymerase involved in DNA repair	COG0389		
LSEI_1081	Superfamily II DNA helicase	COG0514		
LSEI_1103	Transposase and inactivated derivatives	COG3464		
LSEI_1241	Transposase and inactivated derivatives	COG2963		
LSEI_1242	Transposase and inactivated derivatives	COG2801		
LSEI_1295	ATP-dependent exoDNAse (exonuclease V), alpha subunit - helicase superfamily I member	COG0507		
LSEI_1346	Nuclease subunit of the excinuclease complex	COG0322		
LSEI_1368	Site-specific recombinase XerD	COG4974		
LSEI_1415	Site-specific recombinase XerD	COG4974		
LSEI_1535	Transposase and inactivated derivatives	COG3293		
LSEI_1709	5'-3' exonuclease (including N-terminal domain of PolI)	COG0258		
LSEI_1891	Transposase and inactivated derivatives	COG3801		
LSEI_1907	Transposase and inactivated derivatives	COG2801		
LSEI_1916	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_1978	Site-specific recombinase XerD	COG4974		
LSEI_1995	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_2005	Transposase and inactivated derivatives	COG2801		
LSEI_2082	ATP-dependent exoDNAse (exonuclease V), alpha subunit - helicase superfamily I member	COG0507		
LSEI_2091	Site-specific recombinase XerD	COG4974		
LSEI_2181	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_2212	3-methyladenine DNA glycosylase	COG2818		
LSEI_2234	DNA mismatch repair enzyme (predicted ATPase)	COG0323		
LSEI_2235	Mismatch repair ATPase (MutS family)	COG0249		
LSEI_2347	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_2404	Transposase and inactivated derivatives, IS30 family	COG2826		
LSEI_2595	Mg-dependent DNase	COG0084		
LSEI_2637	Excinuclease ATPase subunit	COG0178		
LSEI_2649	RecA/RadA recombinase	COG0468		
LSEI_2877	Transposase and inactivated derivatives, IS30 family	COG2826		
	<b>Restriction-modification system</b>			
LSEI_2094	Type I restriction-modification system methyltransferase subunit	COG0286		
LSEI_2095	Type I site-specific restriction-modification system, R (restriction) subunit and related helicases	COG0610		

**Carbohydrate metabolism**

Locus tag	Name	COG number
<b>ABC-type transporters</b>		
LSEI_0289	ABC-type sugar transport system, periplasmic component	COG1653
LSEI_0307	ABC-type ribose transport system, auxiliary component	COG1869
LSEI_0308	ABC-type sugar transport system, ATPase component	COG1129
LSEI_0988	ABC-type sugar transport system, periplasmic component	COG1653
LSEI_0989	ABC-type sugar transport system, permease component	COG0395
LSEI_0990	ABC-type sugar transport systems, permease components	COG1175
LSEI_2762	ABC-type sugar transport system, periplasmic component	COG1653
LSEI_2765	ABC-type sugar transport systems, ATPase components	COG3839
LSEI_2840	ABC-type sugar transport system, periplasmic component	COG1653
<b>Phosphotransferase systems</b>		
LSEI_0119	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	COG1263
LSEI_0276	Phosphotransferase system cellobiose-specific component IIC	COG1455
LSEI_0324	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)	COG1762
LSEI_0366	Phosphotransferase system, mannose/fructose/N-acetylgalactosamine-specific component IIC	COG3715
LSEI_0370	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	COG1263
LSEI_0374	Phosphotransferase system IIA components	COG2190
LSEI_0402	Phosphotransferase system, mannose/fructose/N-acetylgalactosamine-specific component IIC	COG3715
LSEI_0403	Phosphotransferase system, mannose/fructose/N-acetylgalactosamine-specific component IID	COG3716
LSEI_0449	Phosphotransferase system cellobiose-specific component IIA	COG1447
LSEI_0559	Phosphotransferase system, mannose/fructose-specific component IIA	COG2893
LSEI_0561	Phosphotransferase system, mannose/fructose/N-acetylgalactosamine-specific component IIC	COG3715
LSEI_0631	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	COG1263
LSEI_0672	Phosphotransferase system, galactitol-specific IIB component	COG3414
LSEI_0673	Phosphotransferase system, galactitol-specific IIC component	COG3775
LSEI_1347	Phosphotransferase system, fructose-specific IIC component	COG1299
LSEI_2077	Phosphotransferase system cellobiose-specific component IIC	COG1455
LSEI_2101	Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific	COG1263
LSEI_2247	Phosphotransferase system cellobiose-specific component IIB	COG1440
LSEI_2596	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)	COG1762
LSEI_2666	Phosphotransferase system, mannose/fructose/N-acetylgalactosamine-specific component IIC	COG3715
LSEI_2736	Phosphotransferase system, galactitol-specific IIB component	COG3414
LSEI_2740	Phosphotransferase system, fructose-specific IIC component	COG1299
LSEI_2741	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)	COG1762
LSEI_2815	Phosphotransferase system cellobiose-specific component IIC	COG1455
LSEI_2827	Phosphotransferase system, mannose/fructose/N-acetylgalactosamine-specific component IID	COG3716
LSEI_2829	PTS system mannose/fructose-specific transporter subunit IIA	COG3444
LSEI_A05	Phosphotransferase system cellobiose-specific component IIB	COG1440

**Other transporters**

LSEI_0058	Arabinose efflux permease	COG2814
LSEI_0172	Arabinose efflux permease	COG2814
LSEI_0229	H <sup>+</sup> /gluconate symporter and related permeases	COG2610
LSEI_0429	Arabinose efflux permease	COG2814
LSEI_0578	Na <sup>+</sup> /melibiose symporter and related transporters	COG2211
LSEI_0878	Arabinose efflux permease	COG2814
LSEI_1066	Arabinose efflux permease	COG2814
LSEI_1125	Putative glucose uptake permease	COG4975
LSEI_1140	Arabinose efflux permease	COG2814
LSEI_1283	Arabinose efflux permease	COG2814
LSEI_1758	Arabinose efflux permease	COG2814
LSEI_2188	Arabinose efflux permease	COG2814
LSEI_2211	Permeases of the drug/metabolite transporter (DMT) superfamily	COG0697
LSEI_2230	Arabinose efflux permease	COG2814
LSEI_2274	Arabinose efflux permease	COG2814
LSEI_2277	Arabinose efflux permease	COG2814
LSEI_2342	Arabinose efflux permease	COG2814
LSEI_2416	Arabinose efflux permease	COG2814
LSEI_2749	Arabinose efflux permease	COG2814
<b>Enzymes</b>		
LSEI_0279	Phosphopentomutase	COG1015
LSEI_0291	N-acetyl-beta-hexosaminidase	COG3525
LSEI_0369	Alpha-galactosidases/6-phospho-beta-glucosidases, family 4 of glycosyl hydrolases	COG1486
LSEI_0404	Alpha-glucosidases, family 31 of glycosyl hydrolases	COG1501
LSEI_0406	Glycosidases	COG0366
LSEI_0432	Fructose/tagatose bisphosphate aldolase	COG0191
LSEI_0448	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase	COG2723
LSEI_0564	Beta-fructosidases (levanase/invertase)	COG1621
LSEI_0677	Fructose-1-phosphate kinase and related fructose-6-phosphate kinase (PfkB)	COG1105
LSEI_0700	Beta-glucosidase-related glycosidases	COG1472
LSEI_0821	Predicted sugar phosphatases of the HAD superfamily	COG0647
LSEI_0982	Trehalose and maltose hydrolases (possible phosphorylases)	COG1554
LSEI_1036	Alpha-galactosidase	COG3345
LSEI_1104	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase	COG2723
LSEI_1293	Phosphoglycerate mutase	COG0406
LSEI_1348	Tagatose-6-phosphate kinase	COG1105
LSEI_1365	Pyruvate kinase	COG0469
LSEI_1681	6-phosphogluconate dehydrogenase	COG0362
LSEI_1858	Citrate lyase beta subunit	COG2301
LSEI_1992	Fructose-2,6-bisphosphatase	COG0406
LSEI_2035	Glycosidases	COG0366
LSEI_2036	Glucan phosphorylase	COG0058
LSEI_2037	Glycogen synthase	COG0297
LSEI_2038	ADP-glucose pyrophosphorylase	COG0448
LSEI_2039	ADP-glucose pyrophosphorylase	COG0448
LSEI_2040	1,4-alpha-glucan branching enzyme	COG0296
LSEI_2045	Uncharacterized protein conserved in bacteria	COG3855

LSEI_2049	Capsular polysaccharide biosynthesis protein	COG4464
LSEI_2102	Glycosidases	COG0366
LSEI_2104	Beta-fructosidases (levanase/invertase)	COG1621
LSEI_2134	Phosphoglycerate mutase 1	COG0588
LSEI_2191	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase	COG2723
LSEI_2334	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	COG0574
LSEI_2644	Fructose-2,6-bisphosphatase	COG0406
LSEI_2684	Alpha-galactosidases/6-phospho-beta-glucosidases, family 4 of glycosyl hydrolases	COG1486
LSEI_2731	Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases	COG0235
LSEI_2733	Putative L-xylulose-5-phosphate 3-epimerase	COG3623
LSEI_2769	Alpha-glucosidases, family 31 of glycosyl hydrolases	COG1501
LSEI_A04	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase	COG2723
<b>Regulators</b>		
LSEI_0250	Transcriptional regulator of sugar metabolism	COG1349
LSEI_1659	Transcriptional regulator/sugar kinase	COG1940
LSEI_2734	Transcriptional regulator of sugar metabolism	COG1349

#### Amino acids metabolism

Locus tag	Name	COG number
<b>ABC-type transporters</b>		
LSEI_0065	ABC-type proline/glycine betaine transport systems, permease component	COG1174
LSEI_0066	ABC-type proline/glycine betaine transport systems, ATPase components	COG1125
LSEI_0265	ABC-type branched-chain amino acid transport systems, periplasmic component	COG0683
LSEI_0269	ABC-type branched-chain amino acid transport systems, ATPase component	COG0410
LSEI_0601	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	COG0834
LSEI_1007	ABC-type spermidine/putrescine transport system, permease component I	COG1176
LSEI_1207	ABC-type oligopeptide transport system, periplasmic component	COG4166
LSEI_1344	ABC-type amino acid transport system, permease component	COG0765
LSEI_1608	ABC-type dipeptide transport system, periplasmic component	COG0747
LSEI_1612	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	COG0444
LSEI_1887	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	COG0444
LSEI_1892	ABC-type oligopeptide transport system, periplasmic component	COG4166
LSEI_2065	ABC-type oligopeptide transport system, periplasmic component	COG4166
LSEI_2797	ABC-type oligopeptide transport system, periplasmic component	COG4166
LSEI_2837	ABC-type amino acid transport system, permease component	COG0765
LSEI_2838	ABC-type polar amino acid transport system, ATPase component	COG1126
LSEI_2848	ABC-type polar amino acid transport system, ATPase component	COG1126
LSEI_2849	ABC-type amino acid transport system, permease component	COG0765
LSEI_2850	ABC-type amino acid transport system, permease component	COG0765
LSEI_A10	ABC-type amino acid transport system, permease component	COG0765
LSEI_A11	ABC-type polar amino acid transport system, ATPase component	COG1126
LSEI_A12	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	COG0834
<b>Other transporters</b>		
LSEI_0229	H <sup>+</sup> /gluconate symporter and related permeases	COG2610
LSEI_0303	Amino acid transporters	COG0531
LSEI_0385	Amino acid transporters	COG0531
LSEI_0441	Amino acid transporters	COG0531
LSEI_0471	Amino acid transporters	COG0531
LSEI_0633	Amino acid transporters	COG0531
LSEI_0642	Gamma-aminobutyrate permease and related permeases	COG1113
LSEI_1031	Gamma-aminobutyrate permease and related permeases	COG1113
LSEI_1813	Dipeptide/tripeptide permease	COG3104
LSEI_1996	Amino acid transporters	COG0531
LSEI_2211	Permeases of the drug/metabolite transporter (DMT) superfamily	COG0697
LSEI_2236	Amino acid transporters	COG0531
LSEI_2873	Gamma-aminobutyrate permease and related permeases	COG1113
<b>Peptidases, proteases</b>		
LSEI_0033	Dipeptidase A	COG4690
LSEI_0206	Pyroglutamate-carboxylate peptidase	COG2039
LSEI_0246	Oligoendopeptidase F	COG1164
LSEI_0470	Amino peptidase N	COG0308
LSEI_0745	Oligoendopeptidase F	COG1164
LSEI_1094	Endopeptidase, M10 family	

LSEI_1095	Endopeptidase, M10 family	
LSEI_1472	Oligoendopeptidase O	COG1164
LSEI_1651	Xaa-Pro dipeptidyl-peptidase	COG2936
LSEI_2186	Dipeptidase A	COG4690
LSEI_2329	Aminopeptidase C	COG3579
LSEI_2582	Di/tripeptidase	
LSEI_2745	Exopeptidase, M20(A) family	COG0624
<b>Enzymes related in amino acids pathways</b>		
LSEI_0074	Tryptophan synthase alpha chain	COG0159
LSEI_0075	Tryptophan synthase beta chain	COG0133
LSEI_0077	Indole-3-glycerol phosphate synthase	COG0134
LSEI_0097	Tetrahydrodipicolinate N-succinyltransferase	COG2171
LSEI_0099	Aspartokinases	COG0527
LSEI_0100	Diaminopimelate epimerase	COG0253
LSEI_0332	Predicted ornithine cyclodeaminase, mu-crystallin homolog	COG2423
LSEI_0381	Cystathionine beta-lyases/cystathionine gamma-synthases	COG0626
LSEI_0382	Predicted amino acid racemase	COG3457
Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase]		
LSEI_0433		COG0028
LSEI_0625	5-methyltetrahydropteroylglutamate / homocysteine S-methyltransferase	COG0620
LSEI_0764	Methionine synthase II (cobalamin-independent)	COG0620
LSEI_0865	Aspartate/tyrosine/aromatic aminotransferase	COG0436
Bifunctional PLP-dependent enzyme with beta-cystathionase and maltose regulon repressor activities		
LSEI_0894		COG1168
LSEI_1182	Selenocysteine lyase	COG0520
LSEI_1199	L-serine deaminase	COG1760
LSEI_1289	Cysteine sulfinatase desulfinase/cysteine desulfurase and related enzymes	COG1104
LSEI_1434	ATP phosphoribosyltransferase	COG0040
LSEI_1453	Carbamoylphosphate synthase small subunit	COG0505
LSEI_1554	N-formylglutamate amidohydrolase	COG3741
Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase]		
LSEI_1784		COG0028
LSEI_2041	Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lyase	COG0115
LSEI_2149	Homoserine kinase	COG0083
LSEI_2152	Aspartokinases	COG0527
LSEI_2357	Gamma-glutamyl phosphate reductase	COG0014
LSEI_2556	Glutamate synthase domain 2	COG0069
LSEI_2584	Phosphoribosylpyrophosphate synthetase	COG0462
LSEI_2712	Selenocysteine synthase [seryl-tRNA <sup>Ser</sup> selenium transferase]	COG1921
LSEI_2744	Asparaginase	COG1446
Acetylmethionine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases		
LSEI_2745		COG0624
LSEI_2747	Aspartate/tyrosine/aromatic aminotransferase	COG0436
Acetylmethionine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases		
LSEI_2796		COG0624
LSEI_2855	Transglutaminase-like enzymes, putative cysteine proteases	COG1305

**Transport**

Locus tag	Name	COG number
<b>ABC-type transporters</b>		
LSEI_0295	ABC-type cobalt transport system, ATPase component	COG1122
LSEI_0411	ABC-type metal ion transport system, periplasmic component/surface antigen	COG1464
LSEI_0938	ABC-type phosphate transport system, permease component	COG0581
LSEI_1000	ABC-type Na <sup>+</sup> efflux pump, permease component	COG1668
LSEI_1178	ABC-type metal ion transport system, ATPase component	COG1135
LSEI_2174	ABC-type Na <sup>+</sup> efflux pump, permease component	COG1668
LSEI_2421	ABC-type metal ion transport system, periplasmic component/surface adhesin	COG0803
LSEI_2433	ABC-type metal ion transport system, periplasmic component/surface adhesin	COG0803
LSEI_2445	ABC-type phosphate/phosphonate transport system, periplasmic component	COG3221
LSEI_2446	ABC-type phosphate/phosphonate transport system, ATPase component	COG3638
ABC-type nitrate/sulfonate/bicarbonate transport systems, periplasmic components		
LSEI_2450		COG0715
LSEI_2453	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	COG1116
LSEI_2454	ABC-type nitrate/sulfonate/bicarbonate transport system, permease component	COG0600
<b>Other transporters</b>		
LSEI_0030	Mn <sup>2+</sup> and Fe <sup>2+</sup> transporters of the NRAMP family	COG1914
LSEI_0589	Na <sup>+</sup> /H <sup>+</sup> antiporter NhaD and related arsenite permeases	COG1055
LSEI_0644	NhaP-type Na <sup>+</sup> /H <sup>+</sup> and K <sup>+</sup> /H <sup>+</sup> antiporters	COG0025
LSEI_0707	Cation transport ATPase	COG0474
LSEI_0728	Permease	COG0628
LSEI_0738	Mg <sup>2+</sup> and Co <sup>2+</sup> transporters	COG0598
LSEI_0749	Cation transport ATPase	COG0474
LSEI_0973	Chloride channel protein EriC	COG0038
LSEI_1004	NhaP-type Na <sup>+</sup> /H <sup>+</sup> and K <sup>+</sup> /H <sup>+</sup> antiporters	COG0025
LSEI_1012	Cation transport ATPase	COG0474
LSEI_1284	Uncharacterized protein involved in tellurite resistance	COG3853
LSEI_1658	Rhodanese-related sulfurtransferase	COG0607
LSEI_1822	Fe <sup>2+</sup> transport system protein B	COG0370
LSEI_1987	Kef-type K <sup>+</sup> transport systems, membrane components	COG0475
LSEI_1990	Mn <sup>2+</sup> and Fe <sup>2+</sup> transporters of the NRAMP family	COG1914
LSEI_2281	Cation transport ATPase	COG0474
LSEI_2333	K <sup>+</sup> transporter	COG3158
LSEI_2413	Mn <sup>2+</sup> and Fe <sup>2+</sup> transporters of the NRAMP family	COG1914
LSEI_2414	Cation transport ATPase	COG0474
LSEI_2590	Co/Zn/Cd efflux system component	COG1230
LSEI_2616	Polyphosphate kinase	COG0855
LSEI_2632	Tellurite resistance protein and related permeases	COG1275
LSEI_2794	Na <sup>+</sup> /H <sup>+</sup> antiporter NhaD and related arsenite permeases	COG1055
LSEI_2799	Cation transport ATPase	COG2217
LSEI_A01	Cation transport ATPase	COG2217

**Cell wall**

Locus tag	Name	COG number	Description
LSEI_0158	cell surface protein	N/A	
LSEI_0249	Cell wall-associated hydrolase	N/A	
LSEI_0291	lacto-N-biosidase	COG3525	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_0411	ABC-type metal ion transport system, periplasmic component/surface antigen	COG1464	
LSEI_0416	Predicted oxidoreductase; Myosin-crossreactive antigen ortholog	COG4716	
LSEI_0455	Uncharacterized protein encoded in toxicity protection region of plasmid R478, contains von Willebrand factor (vWF) domain	COG4932	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_0465	Uncharacterized conserved membrane associated protein	COG5271	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_0564	Beta-fructosidase (levanase/invertase)	COG1621	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_1177	ABC-type metal ion transport system, periplasmic component/surface antigen	COG1464	
LSEI_1580	Membrane-associated Zn-dependent protease 1	COG0750	
LSEI_1905	hypothetical protein	N/A	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_2029	Surface antigen	COG0791	Cell wall-associated hydrolases (invasion-associated proteins)
LSEI_2139	Sortase (surface protein transpeptidase)	COG3764	LPXTG-site transpeptidase (sortase) family protein (TIGR01076)
LSEI_2226	Cell wall-associated hydrolase		
LSEI_2270	Subtilisin-like serine proteases	COG1404	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_2366	Predicted outer membrane protein	COG4932	
LSEI_2421	ABC-type metal ion transport system, periplasmic component/surface adhesin	COG0803	
LSEI_2433	ABC-type metal ion transport system, periplasmic component/surface adhesin	COG0803	
LSEI_2437	Possible cell surface protein	COG3846	Type IV secretory pathway, TrbL components
LSEI_2512	Predicted outer membrane protein	COG4932	
LSEI_2537	cell surface protein	N/A	
LSEI_2660	Subtilisin-like serine protease	N/A	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_2686	Alpha-glucosidase, family 31 of glycosyl hydrolase	N/A	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_2769	Alpha-glucosidases, family 31 of glycosyl hydrolases	COG1501	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_2896	Adhesion exoprotein	N/A	LPXTG-motif cell wall anchor domain (TIGR01167)
LSEI_2897	Adhesion exoprotein	N/A	

**Peptidoglycan biosynthesis proteins**

LSEI_0314	Cell division protein FtsI/penicillin-binding protein 2	COG0768	
LSEI_0810	Membrane carboxypeptidase (penicillin-binding protein)	COG0744	
LSEI_1152	Beta-lactamase class C related penicillin binding protein	COG1680	
LSEI_1481	Membrane carboxypeptidase (penicillin-binding protein)( EC:3.4.-, EC:2.4.1.- )	COG0744	
LSEI_1725	Membrane carboxypeptidase (EC:2.3.2.-, EC:2.4.1.129)	COG0744	

Table S3. Tags used to generate STM library

Tag name	sequence (5'→3')
A1	AGCGAGGGATAGCGGTTGGTCGTTGTCGGTGTGGGTGGAT
A2	GTGCGGGAGCGGGGTATGGAGAGTGTGGTTGGTGTITTT
A3	GGTGAGCGAGGTAGAGGGATGGCTGCGTGGTGGGGTT
A4	GGGGCGGTGTGGATAGCTGGTTAGCGGTAGGTCGGTTT
A5	AGCTGGAGCTCGAGAGTGAAGTGGGGGTGCTTTAGCTAT
A6	CACACCCACGACCAAGCACACGAACGCCCTCTCTATCA
A7	GTCGAGGTTTATATGTTGAGGGGGCGGGTATTGTGTGT
A8	GGTGGGGGGTGTCTCGCTTTCGATGTGGTGTGGGTCT
A9	GGGGCGCTAGGTCGGGGAGTGTTCGCGTTGGGTGAT
A10	AGAGCTGTGGAGTCGATGTATGGTGGGAGAGTGAAGT
A11	ATCACGCACCCGATCTCCAACCAACCCTCTCCACACCC
A12	AGGTGGATATAGAGGGTGGTGGGGGAGCTGGTTGGGAT
B1	CGGGTGTGGTAGATAGATCTATCTCGAGATAGCTAGGT
B2	ATCCACCCACACCGACAACGACCAACCGCTATCCCTCGT
B3	AGCCCCCTAGACCGATCCAAATACCGAGCGATCCATACCT
B4	CACACTCCCTCGCGCTACATCGCCGACCGCTCCAGCTCA
B5	AAACCGACTACCGCTAACAGCTATCCACAGCGCCCC
B6	ATCTGGGGAGCGCGATAGGGTGGGGCGTTAGCTGGATCG
B7	ATGTGTCGCGAGGAGGGTGGGGGCTAGTTCGCGTTTT
B8	ACCAACCCCTCCACACCAACACCGCTCTCTCGATACCG
B9	AGACCAAAACCCACATCGAAAGCGAGATCAACCCCAAC
B10	ACAACACCAAAAACCCCAACAAGATCTCACTATCTCA
B11	ACCTCACTCTCCCAACCAATACATCGAGCTCCACAGCTCT
B12	GGTGTGGGAGTGGGGCGATTAGGGCTGT
C1	GGAGAGGGGTTGGTGTAGGGGGGGCTCGATGTTGGGGAT
C2	GGAGGTCGTTGGCTAGCTGGCGGGTGGAGCATTTTGT
C3	AGTGCTATCGTTATGGCTGGGTCTGGGGCGGGTGGTTAT
C4	AACACCATCGAGCTATAGCGAACCGCGACGACCCCTCCCC
C5	ACAAATCTCACAGCTCAACAACAACCGCCCGCGACTCC
C6	AACCATCTAGAAACAACAACGACAAAACGCACTCTCGCC
C7	AGCTTGGTGGGGGATGTATGTGTATGTTGATGAGGCTCT
C8	GTGGGGGGCCGTACGCGCTCGCGGTGCTGTGTAGGGATCG
C9	CTTTGTGTGGGGATCTTGGCGTGTGTGAGAGCGAGTT
C10	ATAGCGTGGGGGCTGGCTGGCGTGGGAGGTGGTGTGG
C11	AGGGTGTAGAGATAGCGGTTGGGTGTGTAGCTGGTGT
C12	ATCCAAACCAAGCTCCCCCTACCACCTCTATATCCACCT
D1	CCAGAGCCAAGCCACCACAACCCATCAACAAAAGTCCCG
D2	GGGGGTGAGGTGGATCTCGATTGGGGGGTGGTTGGCT
D3	GTGGGGGGTCTCGGTATCGGTGTGTGTGTGGTGTCT
D4	ACCGATCGAGAGCAAAACCACTCGCGACACCAACCTCGT
D5	AACTCTCCCTCCAGCCCGCAGAGCTCCATCTCGACCC
D6	AGAGATTGCGTGGCTTTGGTATAGAGAGGTGGTGGCGG
D7	GGTGGTGGCAGGGTGGTGTGGTGGGGGTTGGTGGT
D8	AACCGCACACCGCAACCCCGCC
D9	AGGTGGTCTGGTTATGTTGGGTGGTGGATCTTGGCT
D10	GGGGGGGTGGTGGGTCTAGAGTGTGGGGTGGCGTTT
D11	ACCCACCAACCGCACTCCCCCATCGAGCTCAACACAC
D12	CACCATCACATCGAACCCCGATATCACGACTATCC
E1	GGGTAGAGAGATTGCGTGTCTTGGCGTTTATCGTGGGT
E2	AGCAGCGACTCCACCAAGAGAGCCCGAGCGCCAGCC
E3	CACCAACAACCGATATCATCCGAGCGACCCCTATCGCT
E4	ATATAGTGAATCGCTGGGGTGGGATATTCGTTCTGT
E5	AGTCCCAACAGCTACCACCCGCAATCAACAACCTCCCC
E6	AGGTGGGGGATGTTGAGCGGGCGGGGAGTGTGTGTCG
E7	AGGTGGGTAGGGGTCTATTCTTGGTGTCTATGGGGCGTT
E8	AGAGAGCACCCGCACTACAACAATCTACCGATCTCTCC
E9	ATCAACCACACACAGCACACCCATAACCATATCGCTCC
E10	AACGCAAGATCCACCACCAACAATAAACCAGAACCCCT
E11	TGGGAGTGTGGTTCGCTGGGTCTAGAGCTCTTGGTAGCG
E12	AACCCCACTCAATCACGCAACACCCCTCGACACCTCCCC
F1	AGGCGATAGGGTGGCTCGGATGATATCGGTGTGGTG
F2	GGAGGGAATAGAGTTGGCGTGGATCGTTGGGTGGAT
F3	AAAAGACAGCCATACAACCTCCCTATCCAAACCCCTTA
F4	CGTGGCGTGTCTTGGAGTTCGGGTGTGGTGGGGAT
F5	CCCAACAACACTACCCTGCTCCACCGCAATCTAGCGCACT
F6	AGCGACACAACCCCGAAATAACCCCGCGACAACATCAA
F7	CCCACACTAGCCTACCCACCCACAGATCGACCCGAGCACCA
F8	GTGTGTGAGCTCGATGGGGGAGTGGGGTGTGGGGGTA
F9	CGAACGCGCGCAAGCCCCCAACTATCTCCAATCTCGAC
F10	ACAGCAATCCATAACAACAACCACTACCCATACAACAAG
F11	TGTGTGTGGGTTTGGGGTATGTGTGTTTGGTGTCT
F12	GGCTATGGGTAGGGATGGGTATGTGGTGGTGGTGGTTA

**Table S4. Primers for tagged mutants relative quantification**

Primer name	Sequence (5' -> 3') *
<i>Reverse primer – common to all tags.</i>	
QPVITAGR	GGCGATTAAGTTGGGTAACG
<i>Forward primers – one per tag</i>	
QTAGA1F	GATAGCGGTTGGTCGTTGTC
QTAGA2F	TGGAGAGTGTGGTTGGTGTT
QTAGA3F	GTGAGCGAGGTAGAGGGATG
QTAGA4F	GGTTAGCGGTAGGTCGGTTC
QTAGA5F	CGAGAGTGAGTGAGGGGTGT
QTAGA6F	CACGACCAAGCACACGAA
QTAGA7F	GGGCGGGTATTGTGTGTT
QTAGA9F	CTAGGTCGGGGGAGTTGTTG
QTAGA10F	ATGTATGGGTGGGAGAGTG
QTAGA11F	AACTCAACCCTCTCCACAC
QTAGA12F	AGGTGGATATAGAGGGTGGTAGG
QTAGB1F	GCTTCGGGTGTGGTAGAT
QTAGB2F	GACAACGACCAACCGCTATC
QTAGB3F	AATACCGAGCGATCCATACC
QTAGB4F	GACCGCTCCAGCTCAaag
QTAGB5F	ACCGACTACCGCTAACCCAG
QTAGB6F	GATAGGTTGGGGCGTTAG
QTAGB7F	GGGGCTAGTTCGCGTTTaag
QTAGB8F	CAACACCGCTCTCTCTCGAT
QTAGB9F	ACATCGAAAGCGAGATCACC
QTAGB10F	CCCACAAGATCTCACTATCTCAa
QTAGB11F	CACTCTCCCCACCCATACAT
QTAGB12F	GGGGCGATTAGGGCTGT
QTAGC1F	GGAGAGGGGTTGGTTGAGG
QTAGC2F	GCGGTTGAGCGATTTGTGT
QTAGC3F	GTTATGGCTGGGTCTGGG
QTAGC4F	caagcttAACACCATCGAGCTAT
QTAGC5Fbis	cttACAAATCTCACCAGCTCC
QTAGC6F	CACAAACGACAAAACGCACT
QTAGC7F	CGGGCGATGTATGTGTATGT
QTAGC8F	GGTCGTGTGTAGGGATCG
QTAGC9Fbis	GTTTGTGAGAGCGAGTTaagc
QTAGC10F	GTGCGAGGTGGTTTGTGG
QTAGC11Fbis	gettAGGGTGCTAGAGATAG
QTAGC12F	CTCCCCCTACCACCTCTAT
QTAGD2F	GAGGTGGATCTCGATTGTGG
QTAGD3F	GGTTGTTCGTGTGTTGGTGT
QTAGD4F	GATCGAGAGCAAACCCACTC
QTAGD5Fbis	CAGAGCTCCATCTCGACC
QTAGD6F	TAGAGAGGTGGTGAGCaag
QTAGD7F	TTGCGAGGGTGGTTTGGTG
QTAGD8F	cttAACCGGACACCGCAAC
QTAGD9F	TGGTTATGTTGGGTGGTG
QTAGD10F	GAGTGTGAGGGTGAGCGTTT
QTAGD11F	ACCCACCAACCCGCACT
QTAGD12F	CACCATCACCATCGAACCAC
QTAGE1F	GGGTAGAGAGATTGCGTGCT
QTAGE2F	CCACCACCAAGAGAGCC
QTAGE3F	ACAACCGATATCATCCGAGC

QTAGE4F	GCTTGGGTGTGGGATATTC
QTAGE5F	GCTACCACCCGCAATACCAA
QTAGE6Fbis	CGGGGGAGTGTGTGTGTCG
QTAGE7Fbis	AGGTGGGTAGGGGTCTATTT
QTAGE8F	CCCGCACTAACAACCAATCT
QTAGE9F	AACCACCACACAGCACACC
QTAGE10F	CATAACCAGAACCACCTaagc
QTAGE11F	GTGTGGTTGCGTGGGTTCTA
QTAGE12F	CCACTCAATCACGCAACACC
QTAGF1F	GCTCGGATGATATCGGTTGT
QTAGF2F	GGAGGGAATAGAGTTGGCG
QTAGF3F	AAAGACCAGCCATCAACTCC
QTAGF4F	GGCGTGTCTTGGGAGTTC
QTAGF5F	CCGCTCCCACGCAATCTA
QTAGF6F	gettAGCGACACAACCCC
QTAGF7F	CACAGATCGACCGAGCAC
QTAGF8Fbis	ATGGGGGAGTGCGGGTT
QTAGF9F	TATCTCCAATCTCGACaagc
QTAGF10F	CAGCAATCCATACAAAAACCAC
QTAGF11F	CGGTTTGGGGTATGTGTGT
QTAGF12F	TGGGTAGGGATGGGTATGTG
<i>Primer couple #1 for total quantification of STM (common for all STM)</i>	
QPVICTRL3F	CTGTGGTATGGCGGTAAGT
QPVICTRL3R	GGTTGCTCTGCACACTCAA
<i>Primer couple #2 for total quantification of STM (common for all STM)</i>	
QPVICTRL4F	CAGTCGGGAAACCTGTGCT
QPVICTRL4R	GAGTCAGTGAGCGGGAAGC

\* small letters correspond to sequences in the invariable arm, bold letters sequences

### III. Caractérisation fonctionnelle de l'opéron LSEI\_0219-0221

Le criblage présenté dans l'article précédent a mis en évidence l'implication de 47 gènes différents dans l'établissement de *L. casei* dans l'intestin. La surface cellulaire constitue le premier site de reconnaissance/contact entre l'hôte et la bactérie et est de ce fait très étudiée (voir synthèse bibliographique). Parmi les 47 gènes, 6 gènes codent pour des protéines qui sont impliquées dans la biosynthèse de la paroi cellulaire. Parmi eux le gène LSEI\_0221 codant pour une D-alanyl-D-alanine carboxypeptidase, penicillin-binding protein (PBP0221), présente la particularité d'être en aval de deux gènes LSEI\_0219 et 0220 codant pour des protéines constituant un système à deux composants (TCS). Les mutants pour ces deux gènes présentent aussi la particularité de ne pas être capables de s'établir *in vivo*. De plus, il s'agit du seul TCS des 16 TCS complets existants chez *L. casei* impliqué dans l'établissement *in vivo* d'après notre criblage (au moins un mutant pour le gène codant soit pour le RR soit pour l'HK pour 10 des 16 TCS). Cet opéron putatif présent chez d'autres lactobacilles a, par conséquent, fait l'objet d'une caractérisation plus approfondie.

Le gène LSEI\_0221 est le seul annoté dans le génome comme ayant la fonction de D-ala-D-ala carboxypeptidase. Cette enzyme est normalement nécessaire à la synthèse du peptidoglycane puisqu'elle permet la coupure entre le 4<sup>ième</sup> et le 5<sup>ième</sup> résidu du pentapeptide du peptidoglycane permettant ainsi la transpeptidation entre les différents brins de peptidoglycane. Le pentapeptide est appelé, chez *L. casei*, pentadepsipeptide car le 5<sup>ième</sup> résidu, normalement un D-Ala, est remplacé par un D-Lac conférant ainsi la résistance à la vancomycine. Le fait que la mutation de ce gène ne soit pas létale pour la bactérie suggère soit une complémentation de la fonction par une autre enzyme soit un défaut d'annotation de la PBP0221.

Le TCS codé par les gènes LSEI\_0219-0220 semble avoir un rôle, de par sa localisation génétique en amont du gène LSEI\_0221, dans la biosynthèse de la paroi cellulaire.

L'article ci-après, en cours de préparation, présente les différents résultats de caractérisation de l'opéron. Des résultats moins aboutis et demandant des répétitions biologiques sont présentés à la suite et en français.

#### Principaux résultats de l'article 4.

Les principaux résultats indiquent que l'insertion du transposon dans un gène de l'opéron a un effet polaire sur les gènes de l'opéron situé en aval. En effet, lorsque le gène LSEI\_0219 est interrompu, la transcription des gènes LSEI\_0220 et 0221 est diminuée ; de même l'interruption du gène LSEI\_0220 génère une diminution du taux de transcrits du gène LSEI\_0221. L'incapacité des trois mutants à s'établir *in vivo* peut être causée par la simple inactivation ou la diminution du taux de transcrits du gène LSEI\_0221.

La fonctionnalité du TCS a été vérifiée et les transcrits des gènes impliqués dans la synthèse du peptidoglycane ont été quantifiés chez les mutants par rapport à la souche parentale. Le TCS semble être impliqué dans la régulation de 17 des 50 gènes testés.

La diminution ou l'absence de transcrits du gène LSEI\_0221 engendre une différence de composition du peptidoglycane par rapport à la souche parentale. Cette différence indique que le gène LSEI\_0221 code plutôt pour une protéine ayant la fonction de L,D-carboxypeptidase que de D-Ala-D-Ala carboxypeptidase.

Une diminution de présentation de deux protéines annotées comme antigènes de surface ainsi qu'une augmentation de la sensibilité aux peptides antimicrobiens (hBD1, hBD2 et CCL20) ont été observées chez les trois mutants.

## The LSEI\_0219-0221 operon: a central interface of the surface architecture of *Lactobacillus casei* that conditions its resistance to the gut.

Hélène Scornec<sup>1</sup>, Thierry Pédrón<sup>2, 3</sup>, Richard Wheeler<sup>4, 5</sup>, Ivo Gompert Boneca<sup>4, 5</sup>, Philippe J. Sansonetti<sup>2, 3, 6</sup>, Jean-François Cavin<sup>1</sup>, Hélène Licandro-Seraut<sup>1</sup>.

<sup>1</sup>UMR A PAM, AgroSup Dijon / Université de Bourgogne, 1 Esplanade Erasme, 21000, Dijon, France

<sup>2</sup>Unité de Pathogénie Microbienne Moléculaire, Institut Pasteur, 28 Rue du Docteur Roux, 75724, Paris cedex 15, France

<sup>3</sup>Unité INSERM 786, Institut Pasteur, Paris, France

<sup>4</sup>Institut Pasteur, Unité de Biologie et Génétique de la Paroi Bactérienne, Paris 75015, France

<sup>5</sup>INSERM, Avenir group, Paris 75015, France

<sup>6</sup>Chaire de Microbiologie et Maladies Infectieuses, Collège de France, 11 Place Marcelin Berthelot, 75005, Paris, France

### **Abstract**

The cell surface is the primary recognition/contact site between the bacterium and the host. Among the 47 genes previously identified as required for the establishment of *L. casei* in the gut, we focused on three genes, LSEI\_0219, 0220 and 0221, organized as an operon. LSEI\_0219 and LSEI\_0220 encode a two-component system (TCS) and LSEI\_0221 encodes a d-Alanyl-d-Alanine carboxypeptidase which is a peptidoglycan biosynthesis enzyme. Also, we explored the physiological role of these three genes, the impact of their inactivation on the bacterial cell wall architecture and on the consequences in the frame of bacterial survival in the gut. The TCS, which has been functionally validated by biochemistry analysis, is involved in the transcriptional regulation of several genes in the peptidoglycan biosynthesis. The analysis of the muropeptides composition of peptidoglycan of mutants showed that the protein encoded by gene LSEI\_0221 has an L,D-carboxypeptidase activity. The absence of this protein causes a decrease of the quantity of two surface antigens, LSEI\_0020 corresponding to p40 or msp2 whose implication in the host epithelial homeostasis as been recently studied (Yan *et al.*, J. Biol. Chem. 2013, 288:30742-30751) and LSEI\_2029 which has never been characterized before. The inactivation of each of these three genes induces a susceptibility to antimicrobial peptides (hBD1, hBD2 and CCL20) which could be a main cause of the gut establishment deficiency. We showed that this operon is necessary for antigens presentation and a suitable architecture of the bacterial surface to establish in the gut and resist to host defenses.

### **Introduction**

As natural inhabitants of fermented food and feed (plants, meat, dairy) and also oral cavity, gut and genital tracts, *Lactobacillus* species share a good genetic arsenal to fit new, sometimes harsh, environments/niches (1). Their high adaptability to stresses results from an accurate coordination of cellular processes (production of chaperones and DNA repair proteins, induction of metabolic pathways or transport systems, modifications of membrane composition) mediated by networks of regulators and also two-component systems (2). Their resistance can also be attributed to their cell wall architecture which is the base for the maintenance of cell shape and integrity and, via the proteins exposed, for direct interaction with biotic or abiotic environment (3). The cell wall of *Lactobacillus* is composed of a peptidoglycan layer decorated with teichoic acids and proteins like peptidoglycan hydrolases, LPxTG proteins...that surround the cytoplasmic membrane.

As concern *Lactobacillus* adaptation to the gut, we have recently identified a core of 47 required genes in *Lactobacillus casei* essential for gut establishment, using signature tagged mutagenesis coupled with a screening in ligated rabbit ileal-loop (4). Among them, three consecutive, identically-oriented genes: LSEI\_0219 and LSEI\_0220, and LSEI\_0221. *Lactobacillus casei* is one of the best equipped Lactic Acid Bacteria (LAB) to sense and respond to environmental changes since its genome possesses 16 complete and one incomplete two-component systems and 124 regulators. The LSEI\_0219-0220 genes caught our attention since they encode the only TCS required for gut establishment (4). LSEI\_0221 is the only gene annotated as a putative D-ala-D-ala carboxypeptidase, penicillin-binding protein in *L. casei*, according to genome annotation (5). D-ala-D-ala carboxypeptidase is a hydrolase involved in peptidoglycan maturation which cleaves the pentapeptide side chain of peptidoglycan stand between the fourth and the fifth (and last) alanine residues thereby allowing transpeptidation of two peptidoglycan strands. For *L. casei*, the last D-ala residue is substituted by a D-lactate, conferring vancomycin resistance (6). Genetic location of LSEI\_0219-0220 presumes a role of this TCS in the cell wall biogenesis.

In this study, we explored the physiological role of LSEI\_0219, 0220 and 0221 genes, the impact of their inactivation on the bacterial cell wall architecture and the consequences in the frame of bacterial survival in the gut.

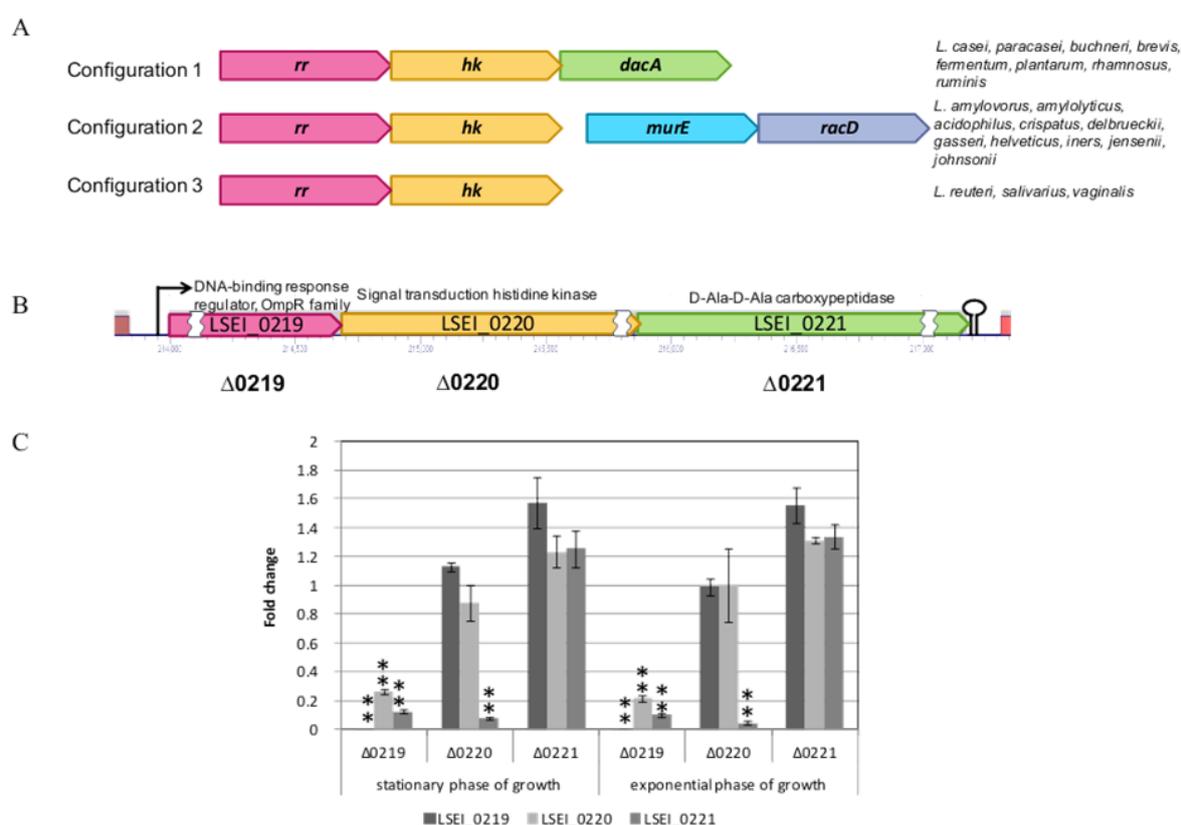
## **Results**

### **Predictive analysis of operon LSEI\_0219-0221 and polar effect of the transposon insertion**

The LSEI\_0219-0221 predicted operon is composed of three genes (Fig. 1A). The LSEI\_0219 gene belongs to the OmpR Family and is predicted to act as a response regulator (RR), named RR0219 in this study; the LSEI\_0220 gene is predicted to act as a signal transduction histidine kinase (HK), named HK0220; the LSEI\_0221 gene is predicted to act as a D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein) named PBP0221.

This type of two-component system (TCS) is largely encountered in the bacterial world. In lactobacilli, it can be found in three different genetic organizations (Fig. 1B). It can be followed by a D-Ala-D-Ala carboxypeptidase, in *Lactobacillus casei*, *paracasei*, *buchneri*, *brevis*, *fermentum*, *plantarum*, *rhamnosus* and *ruminis* for example, while in *Lactobacillus amylovorus*, *amylolyticus*, *acidophilus*, *crispatus*, *delbrueckii bulgaricus*, *delbrueckii lactis*, *gasseri*, *helveticus*, *iners*, *jensenii*, *johnsonii* these TCS genes is directly followed by another putative operon composed of *murE* (UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase) and *racD* (aspartate racemase). In a few species such as *Lactobacillus reuteri*, *salivarius*, *vaginalis* the locus is only composed of the TCS predicted operon. The configurations 1 and 3 (Fig. 1) are observed in vancomycin-resistant lactobacillus species (7-10) except for *L. fermentum* which displays a variable strain-dependent vancomycin-resistance (11). The second configuration is only observed in the vancomycin-susceptible species

LSEI\_0219, 0220 and 0221 are very close since only five nucleotides are between the coding sequence of LSEI\_0219 and 0220 and even LSEI\_0220 and 0221 partially overlap. Thus, we assume that these three genes are co-transcribed, which was confirmed by RT-PCR (data not shown). To determine if transposon insertion in the operon triggers a polar effect, qRT-PCR was performed for the three genes in the three mutants in exponential and stationary phases of growth (Fig. 1C). Transcription was affected downstream of the transposon insertion since transcripts corresponding to LSEI\_0220 and 0221 in mutant  $\Delta 0219$  and transcripts corresponding to LSEI\_0221 in mutant  $\Delta 0220$  were at least 5-fold lower than in the ATCC 334 strain. Thus, one possible assumption is that the defect of these three mutants in gut establishment could result from the inactivation or the transcriptional decrease of LSEI\_0221 only. Also, transcript levels corresponding to genes upstream of the transposon insertion in these three mutants are similar to the ATCC 334, meaning that transposon insertion does not affect the stability of the transcripts upstream.



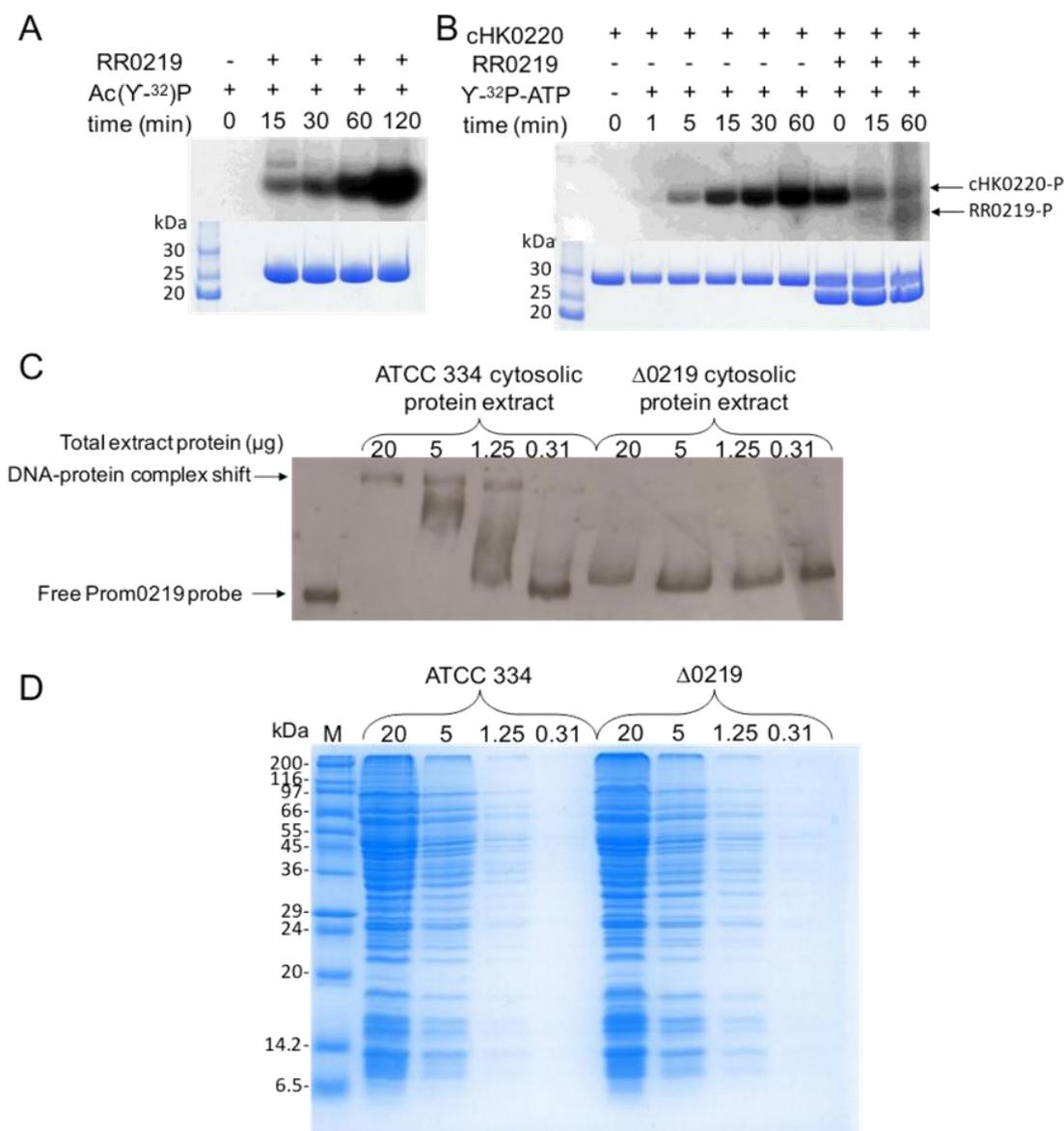
**Fig. 1:** Locus LSEI\_0219-0221 A. Comparison of the organization in different *Lactobacillus* species; *rr*: response regulator gene; *hk*: histidine kinase gene. B, Location of the different transposon insertion in the three mutants  $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$ ; C, Relative transcript levels of *L. casei* mutants in stationary and exponential phase of growth in MRS broth. Transcript levels of each gene are expressed as the relative fold change, with *L. casei* ATCC 334 as the reference condition (fold change = 1). Three biological repeats were performed, and bars indicate standard deviations. Statistical analysis was performed using the unpaired Student *t* test: \*,  $P < 0.01$ ; \*\*,  $P < 0.001$ .

**Functionality of the TCS *in vitro***

The LSEI\_0219 gene was overexpressed in *E. coli* and the resulting RR0219 protein was purified. RR0219 size analyzed by SDS-PAGE was consistent with that expected (27 kDa) (Fig. 2A). The purified RR0219 was able to autophosphorylate using  $^{32}\text{P}$  radiolabeled acetyl-phosphate as the phosphate donor (Fig. 2A).

The cytosolic domain of HK0220, named cHK0220, corresponding to residues 137-396 of the HK0220 protein was overexpressed in *E. coli*, purified and a 29 kDa protein was obtained. It was also able to autophosphorylate using [ $\gamma$ - $^{32}\text{P}$ ] ATP as the phosphate donor which confirms the predicted kinase function. Then the radiolabeled phosphorylated cHK0220 was incubated with the purified RR0219. A partial dephosphorylation of the cHK0220 and a phosphorylation of the RR0219 were visible, demonstrating that the cHK0220 was able to transfer phosphate to the RR0219 (Fig. 2B). The binding phosphate residues are, according the *in silico* analysis, His163 for HK0220 and Asp52 for RR0219.

The phosphorylated or unphosphorylated forms of the purified RR0219 protein were incubated with a 443 bp probe, named Prom0219, corresponding to the whole LSEI\_0218-LSEI\_0219 intergenic region for EMSA. Despite all our attempts to change the experimental conditions (binding buffer, freshly-prepared RR0219, phosphorylated or unphosphorylated, ...), no interaction between the Prom0219 probe and the purified RR could be observed (data not shown). As the failure of EMSA with purified RR0219 could be due to misfolding of RR0219 during protein overexpression procedure, EMSAs were carried out using crude cytosolic protein extracts of *L. casei* ATCC 334 and  $\Delta$ 0219 (Fig. 2C). The Prom0219 probe was shifted by the ATCC 334 protein extract whereas, for the same protein concentration (Fig. 2D), no shift appeared with the protein extract from  $\Delta$ 0219, which does not contain RR0219 and which partly lacks HK0220 and PBP0221. According to these results, the transcriptional regulator of this operon is either RR0219 which needs some cellular elements and/or a protein co-factor to be active at this DNA site, or another transcriptional regulator that would be absent or inactive in the  $\Delta$ 0219 mutant.

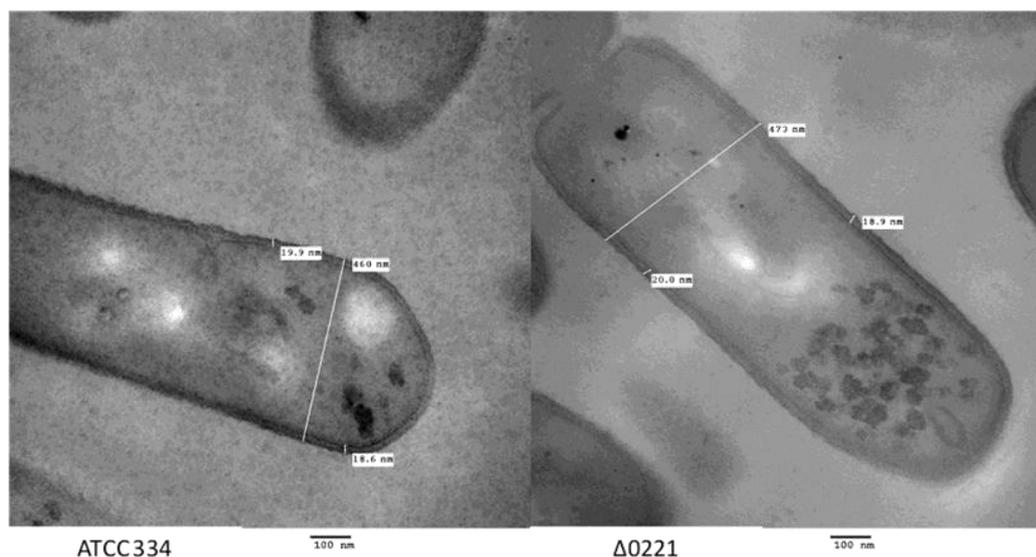


**Fig. 2:** *In vitro* characterizations of the TCS. A: Radiolabeling of RR0219 using [ $\gamma$ - $^{32}$ P] acetyl phosphate after different times of incubation B: Autophosphorylation of the cytosolic domain of HK0220 (cHK0220) when incubated with [ $\gamma$ - $^{32}$ P]ATP. After an incubation of 60 min, the RR0219 protein was added and the phosphotransfer from radiolabeled cHK0220 (cHK0220-P) to RR0219 was observed after different times of co-incubation. C: EMSA of LSEI\_0219 promoter with cellular extract of *L. casei* ATCC 334 and the  $\Delta$ 0219 mutant. D: Corresponding SDS-PAGE 12.5%

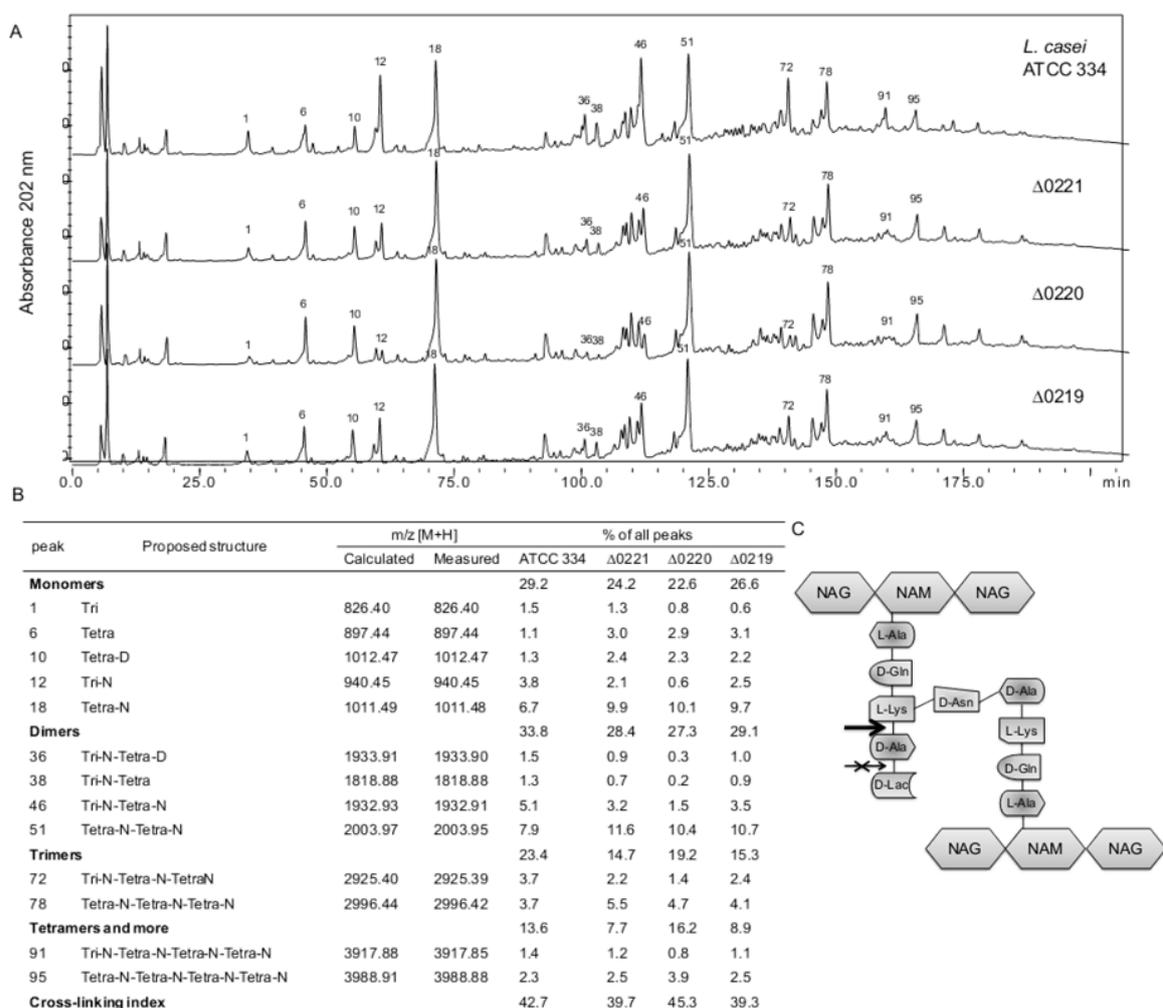
### Physiological role of d-ala-d-ala carboxypeptidase for *L. casei*

The LSEI\_0221 gene encodes a D-Alanyl-D-Alanine carboxypeptidase. This function is essential for peptidoglycan synthesis and, since no other *L. casei* genes are predicted to encode this enzyme, LSEI\_0221 inactivation should be lethal. As a matter of fact,  $\Delta 0221$  is viable and grows as well as the ATCC 334 in culture medium (4). Yet, we wondered if LSEI\_0221 encoded the predicted function. Transmission Electron Microscopy (TEM) analysis of the  $\Delta 0221$  mutant cultivated in exponential and early stationary phases of growth demonstrated that its cell morphology was similar to the ATCC 334 with length between 1 and 2  $\mu\text{m}$  and width between 0.4 and 0.6  $\mu\text{m}$  (Fig. 3). Moreover, its peptidoglycan thickness was not altered in the mutant with value of  $19.4 \pm 3.6$  nm.

Analysis of the muropeptide profiles revealed that the peptidoglycan composition of the mutants  $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$  differs from that of the ATCC 334: tripeptides (peaks 1, 4, 5, 6, 7, 9) are less represented while tetrapeptides (peaks 2, 3, 8 and 10) are more represented than the ATCC 334. However peptidoglycan cross-linking index are not significantly modified in the mutants (Fig. 4A and B). These results suggest a L,D-carboxypeptidase activity, which cleaves the peptide chain of peptidoglycan between L-Lys and D-Ala, rather than the predicted D,D-carboxypeptidase activity (Fig. 4C). Also, the recruitment of other enzymes to compensate the lack of LSEI\_0221 cannot be excluded. Inactivation of LSEI\_0219-0221 operon triggers a significant modification of peptidoglycan composition that could impact other surface components.



**Fig. 3:** Transmission electron microscopy micrographs of *L. casei* ATCC 334 and  $\Delta 0221$  mutant grown in MRS medium until the early stationary phase of growth.



**Fig. 4:** Muropeptides profiles of *L. casei* ATCC 334 and  $\Delta 0219$ ,  $\Delta 0220$  and  $\Delta 0221$  mutants. A. RP-HPLC separation profiles of muropeptides of *L. casei* and mutants, peptidoglycans were digested by mutanolysin, Numbers correspond to the main muropeptides which amounts changed between ATCC 334 and mutants. B. Main muropeptides compositions, Tri, disaccharide tripeptide (L-Ala-D-iGln-L-Lys); Tetra, disaccharide tetrapeptide (L-Ala-D-iGln-L-Lys-D-Ala); Disaccharide, GlcNAc-MurNAc; iGln, isoglutamine; N, D-Asn; D, D-Asp. Percentage of each peak was calculated as the ratio of the peak area to all peak areas. Cross-linking index was calculated according to (12) with the formula  $(1/2 \sum \text{dimers} + 2/3 \sum \text{trimers} + 3/4 \sum \text{tetramers and more}) / \sum \text{all muropeptides}$ . C. Schematic structure of *L. casei* ATCC 334 peptidoglycan. The arrow indicates the cleavage site of PBP0221 coded by LSEI\_0221 gene and the crossed arrow indicates the incorrect, predicted function.

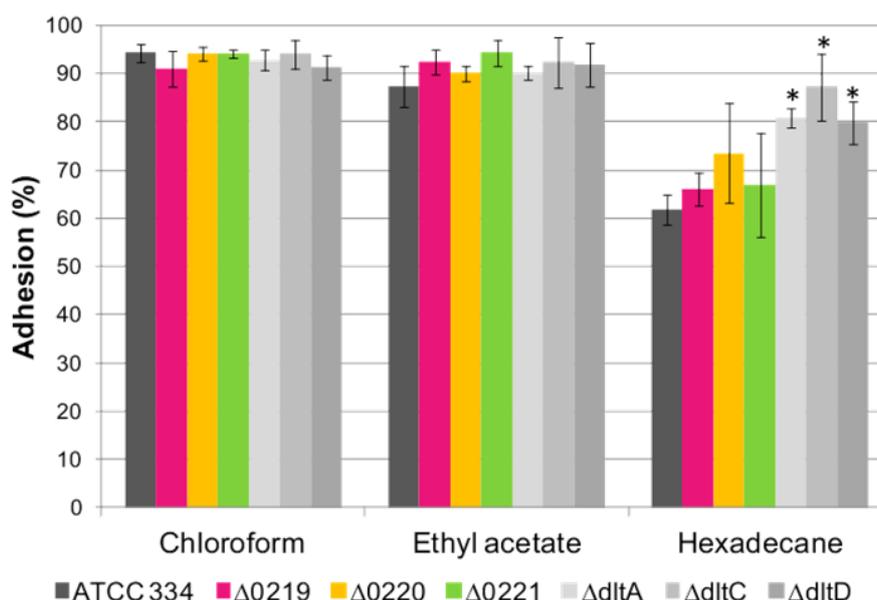
### Consequences of the peptidoglycan composition on the surface properties

The Minimal Inhibitory Concentration (MIC) of four antimicrobials (penicillin, bacitracin, nisine and vancomycin), which inhibit peptidoglycan biogenesis at different levels, was determined for the three mutants and compared to the ATCC 334 (Tab. 1). Concerning vancomycin, mutations in the LSEI\_0219-0221 do not altered peptidoglycan synthesis in a way that would render *L. casei* sensitive to vancomycin, meaning that peptidoglycans of mutants still contains pentadepsipeptides ending with D-Ala-D-Lac. On the contrary, the three mutants displayed an increased sensitivity to penicillin, bacitracin and nisin. An increase of susceptibility to penicillin corresponds to the expected phenotype for penicillin-binding protein defective mutants. The increased bacitracin susceptibility, antibiotic that inhibits undecaprenyl phosphate regeneration, an indispensable step to synthesize a new peptidoglycan strain, suggests a slower renewal of peptidoglycan. Also, mutants are a little more susceptible to nisin than ATCC 334. Nisin seems to go easier through the cell wall, most likely due to a change in surface polarity and disturb the membrane of the mutants.

To assess if mutations trigger a change in charge and hydrophobicity of the bacterial surface, the affinity to chloroform (acidic solvent and electron acceptor), ethyl acetate (basic solvent and electron donor), and hexadecane (hydrophobic solvent) was determined by Microbial Adhesion To Solvents (MATS) tests. In addition to the three mutants of the LSEI\_0219-0221 operon and the ATCC 334, three mutants of the *dlt* operon (*dltA*, *dltC* and *dltD* respectively LSEI\_0794, LSEI\_0796 and LSEI\_0797) were tested because they are altered in teichoic acid D-alanylation, (Fig. 5). The MATS values indicate that *L. casei* ATCC 334 surface displays basic, acidic and hydrophobic properties in agreement with previous observations (13). *dltA*, *dltC* and *dltD* mutants, display a higher surface hydrophobicity than the ATCC 334. Surface properties of LSEI\_0219-0221 mutants are the same as the ATCC 334. Thus, peptidoglycan modifications in mutants do not impact surface properties and do not seem to alter teichoic acid presentation at the cell surface.

**Table 1:** Minimal inhibitory concentration (MIC) of ATCC 334 and three mutants of *L. casei* in  $\mu\text{g.ml}^{-1}$

	Penicillin		Bacitracin			Nisin			Vancomycin		
ATCC 334	0.1	< MIC <	0.2	200	< MIC <	400	10	< MIC <	20	500	< MIC <
$\Delta$ 0219	0.025	< MIC <	0.05	50	< MIC <	100	5	< MIC <	10	500	< MIC <
$\Delta$ 0220	0.05	< MIC <	0.01	50	< MIC <	100	5	< MIC <	10	500	< MIC <
$\Delta$ 0221	0.05	< MIC <	0.01	50	< MIC <	100	5	< MIC <	10	500	< MIC <



**Fig. 5:** MATS tests of ATCC 334 and the six mutants of *L. casei* ( $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$ ,  $\Delta dltA$ ,  $\Delta dltC$ , and  $\Delta dltD$  mutants). Errors bars represent standard deviations between three biological repeats.

### Consequences of the genes inactivation on the regulation of genes involved in the peptidoglycan synthesis

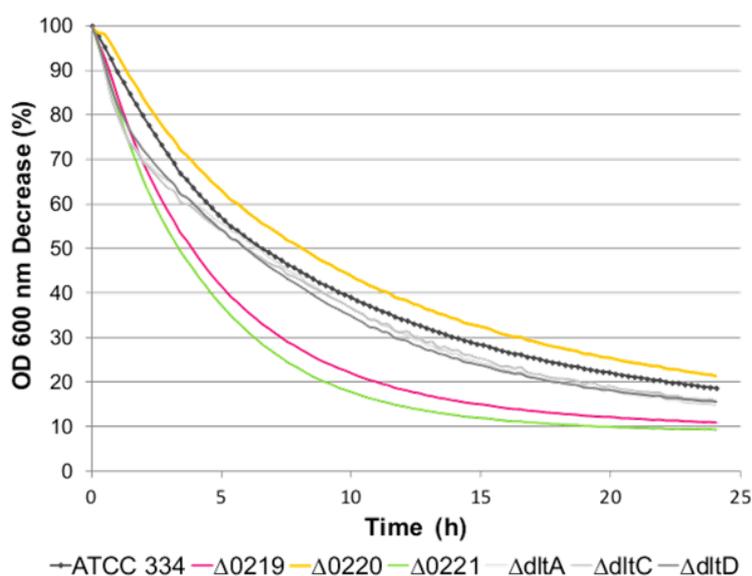
Autolysis profiles of *L. casei* ATCC 334 and the  $\Delta 0219$ ,  $\Delta 0220$  and  $\Delta 0221$  mutants were made by exposing bacteria cells in exponential growth phase ( $OD_{600nm}$  of 0.8) to 0.05% Triton X-100 (v/v) and measuring the  $OD_{600nm}$  during incubation (Fig. 6). Triton X-100 is a detergent known to remove the lipoteichoic acid, which acts as inhibitors on the general bacterial autolytic system (endogenous autolysins) (14). As *L. rhamnosus* GG  $\Delta dltD$  mutant has been reported to d an increased autolysis (15), *L. casei*  $\Delta dltA$ ,  $\Delta dltC$  and  $\Delta dltD$  mutants were also submitted to this autolysis test for comparison. The  $\Delta 0219$  and  $\Delta 0221$  mutants exhibited a higher Triton X-100 induced autolysis rate than the ATCC 334 strain. Autolysis rates of  $\Delta 0219$  and  $\Delta 0221$  mutants were similar to the  $\Delta dltA$ ,  $\Delta dltC$  and  $\Delta dltD$  mutants during the two firsts hours, then they were higher which suggest either a higher recruitment of autolysins in these mutants or a weakened peptidoglycan (PG). Unexpectedly, the  $\Delta 0220$  mutant had an equal autolysis rate than the ATCC 334.

Transcriptomic analysis of PG-related genes in  $\Delta 0219$ ,  $\Delta 0220$  and  $\Delta 0221$  mutants compared to the ATCC 334 was carried out in exponential and stationary phases of growth for two main purposes: determining if the lack of PBP0221 was partly compensated by another enzyme and identifying the possible PG-related genes regulated by the RR0219-HK0220 TCS (Fig. 7 and S1). Two genes are up-regulated in  $\Delta 0221$  mutant, at least in one of the two phases of growth: LSEI\_0393, a 1,4-beta-N-acetylmuramidase (for exponential phase only), and LSEI\_0539, a cell wall-associated hydrolase with no biochemical function attributed. These proteins potentially bring functional compensation of the PBP0221 absence. Overall, whereas only a few PG-related genes are differentially transcribed in exponential phase, a lot are down-regulated in stationary phase of growth in  $\Delta 0219$  mutant and

even more in  $\Delta 0221$ , which suggests a gene regulation directed to slow down the turnover of peptidoglycan. The strongest down-regulated genes in all mutants are LSEI\_0020, 2029 and 2226. LSEI\_2029 and LSEI\_2226 encode cell-wall associated hydrolases. Also, LSEI\_0020 and LSEI\_2029 are annotated as a surface antigen.

By searching differences of relative transcript level between  $\Delta 0219$  (inactivated TCS and lack of LSEI\_0221) and  $\Delta 0221$  (functional TCS and inactivated LSEI\_0221), we proposed 17 candidate genes to be regulated by the RR0219-HK0220 TCS. They can be split into 2 sub-groups: those which were differently expressed in  $\Delta 0221$  (LSEI\_0539, 0810, 0906, 1089, 1269-1272, 1314, 1802 and 2570), meaning a LSEI\_0221-dependent regulation, and those differently expressed in  $\Delta 0219$  (LSEI\_0223, 0251 and 0435), meaning a LSEI\_0221-independent regulation. Also, LSEI\_0393 and 1909 belong to the genes possibly regulated by the TCS but, since their expression profiles vary between exponential and stationary phases, they cannot be classified in these subgroups. All the other tested genes are regulated in a RR0219-HK0220 TCS-independent manner.

Among the tested genes, 23 were transcribed at similar level whatever the condition (Fig. S1 and Table S1)



**Fig. 6:** Autolysis profiles of the ATCC 334 and three mutants of *L. casei*. Exponential culture of strains *L. casei* ATCC 334,  $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$ ,  $\Delta dltA$ ,  $\Delta dltC$ ,  $\Delta dltD$  were grown to an  $OD_{600nm}$  of 0.8; then the cells were pelleted and resuspended in PBS containing 0.05% Triton X-100 (v/v) and incubated at 30°C.

	exponential phase			stationary phase		
	$\Delta 0219$	$\Delta 0220$	$\Delta 0221$	$\Delta 0219$	$\Delta 0220$	$\Delta 0221$
LSEI_0020	-1.30	-1.21	-1.34	<b>-2.52</b>	<b>-2.75</b>	<b>-5.34</b>
LSEI_0249	-2.06	<b>-2.24</b>	-1.66	<b>-2.23</b>	<b>-2.60</b>	-2.31
LSEI_0599	1.03	1.23	1.24	<b>-2.25</b>	-1.72	<b>-2.53</b>
LSEI_1072	-1.38	-1.36	-1.21	<b>-2.07</b>	-1.90	<b>-2.41</b>
LSEI_1116	-1.77	-1.52	-1.30	<b>-2.37</b>	-1.85	<b>-2.39</b>
LSEI_1481	-1.74	-1.81	-1.20	<b>-2.50</b>	-2.05	<b>-2.51</b>
LSEI_1662	-2.12	-1.66	-1.23	<b>-2.17</b>	-1.60	<b>-2.03</b>
LSEI_1725	-1.33	-1.28	1.14	<b>-2.24</b>	-1.77	<b>-2.45</b>
LSEI_2029	-1.44	-1.75	-1.51	<b>-2.90</b>	<b>-4.28</b>	<b>-5.02</b>
LSEI_2226	<b>-2.29</b>	-1.88	-1.73	<b>-2.76</b>	<b>-3.87</b>	<b>-3.98</b>
LSEI_0539	1.02	1.16	<b>4.20</b>	-1.32	-1.42	<b>2.93</b>
LSEI_0810	-1.27	-1.14	-1.23	-1.73	-1.62	<b>-2.25</b>
LSEI_0906	-1.54	-1.24	-1.24	-2.01	-1.67	<b>-2.74</b>
LSEI_1089	-1.21	1.06	-1.45	-1.99	-1.81	<b>-2.61</b>
LSEI_1269	-1.33	-1.13	-1.24	-1.72	-1.48	<b>-2.05</b>
LSEI_1270	-1.34	-1.14	-1.20	-1.81	-1.50	<b>-2.19</b>
LSEI_1271	-1.22	-1.18	-1.11	-1.74	-1.43	<b>-2.31</b>
LSEI_1272	-1.17	-1.35	-1.25	-1.61	-1.40	<b>-2.17</b>
LSEI_1314	1.10	1.35	-1.22	-1.62	-1.52	<b>-2.42</b>
LSEI_1802	-1.07	-1.15	1.09	-1.65	-1.61	<b>-2.06</b>
LSEI_2550	-1.24	-1.15	-2.81	-1.45	-1.18	<b>-3.76</b>
LSEI_2570	-1.16	-1.26	-1.09	-1.68	-1.54	<b>-2.14</b>
LSEI_0393	1.10	1.01	<b>2.24</b>	<b>-2.02</b>	-1.84	1.22
LSEI_1909	<b>-2.82</b>	-2.23	-1.89	-1.17	-1.03	<b>-2.58</b>
LSEI_0223	-1.33	1.22	-1.07	<b>-2.15</b>	-1.24	-1.84
LSEI_0251	-1.72	1.42	1.21	<b>-2.32</b>	-1.10	-1.24
LSEI_0435	<b>-2.66</b>	<b>-2.44</b>	-1.18	1.04	1.17	1.08

Adaptation or compensation of the lack of LSEI\_0221.  
TCS-independent regulation.

Adaptation or compensation of the lack of LSEI\_0221.  
Possible TCS-dependent regulation.

Possible TCS-dependent regulation.  
No link with the lack of LSEI\_0221.

}

**Fig. 7:** Relative transcript levels of genes implicated in peptidoglycan synthesis in  $\Delta 0219$ ,  $\Delta 0220$  and  $\Delta 0221$  mutants compared to the ATCC 334 grown either exponential or stationary phase. Values are the mean RTL obtained for three biological repeats. In bold letters, significant values; in red, up-regulation,  $p < 0.01$ ; in light red, up-regulation,  $p < 0.05$ ; in blue, down-regulation,  $p < 0.01$ ; in light blue, down-regulation,  $p < 0.05$ . Genes for which no significant value was obtained are not presented. The list of the tested genes and their corresponding annotations are presented as supplemental information.

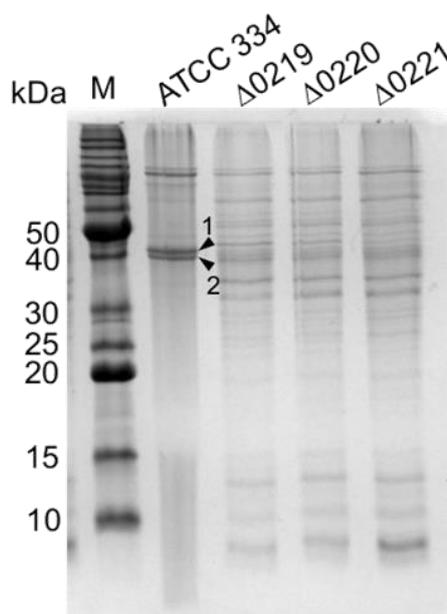
### Impact of the gene inactivation on the surface proteins

Surface protein isolations by LiCl extraction were carried out for the ATCC 334,  $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$  and analyzed by SDS-PAGE (Fig. 8). In the mutants' profiles, two bands, corresponding to proteins of about 40 kDa, were 5-fold lower than the ATCC 334 (Fig. 8). Proteins sequence analysis of these spots revealed that they correspond to LSEI\_0020 and LSEI\_2029 genes, both annotated as surface antigens. This decrease is consistent with the down-regulation of LSEI\_0020 and LSEI\_2029 genes in mutants (see above). The protein encoded by LSEI\_0020 is the orthologous of Msp2 (or named p40) of *L. rhamnosus* GG (16) and *L. casei* BL23 (85% and 95% identities respectively). The N-terminal part of LSEI\_2029 protein sequence presented 46% identity with the

224 first amino acids of LSEI\_0020 whereas the C-terminal part presented 50% identity the 112 amino acids of the C-terminal part of LSEI\_0281, the Msp1 (p75) orthologous.

The very low level of these two proteins in mutants could result from an alteration of their peptidoglycan anchoring (because of the change in PG composition) and/or from a transcriptional regulation. The differential expression of LSEI\_2029 suggests at least the second hypothesis. This regulation would be RR0219-HK0220 TCS-independent.

The other differences in protein profiles between ATCC 334 and mutants should be due to intracellular proteins which should be more released from mutants because of their increased susceptibility to the extraction procedure.



**Fig. 8:** SDS-PAGE of surface proteins extracts from *L. casei* ATCC 334 and  $\Delta 0219$ ,  $\Delta 0220$  and  $\Delta 0221$  mutants in stationary phase of growth. 1: Surface antigen (LSEI\_2029), 2: Surface antigen (LSEI\_0020)

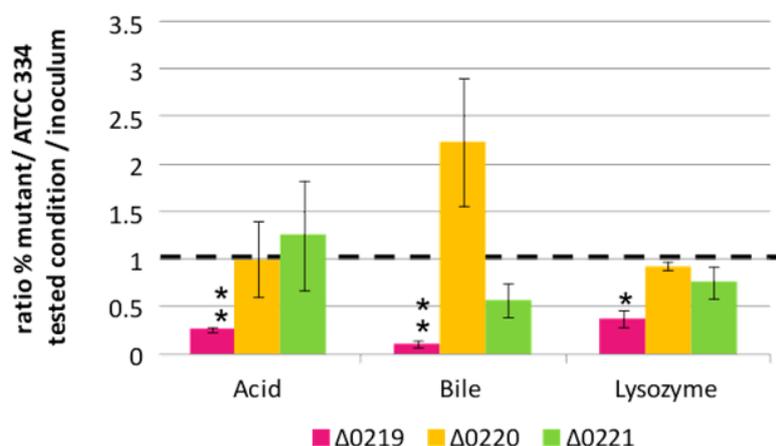
### Impact of gene inactivation on some deleterious factors characterizing the gut

Adhesion to eukaryotic components is one of the bacterial strategies to escape gut stresses. Thus, *L. casei* ATCC 334 and  $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$  mutants were tested for their ability to adhere to polystyrene coated with mucin. The adhesion to mucin was similar for the ATCC 334 and the mutants (about 1.5%) (data not shown) . Thus the change in PG composition and the decrease of surface antigens did not impact the mucin adhesion ability of *L. casei*.

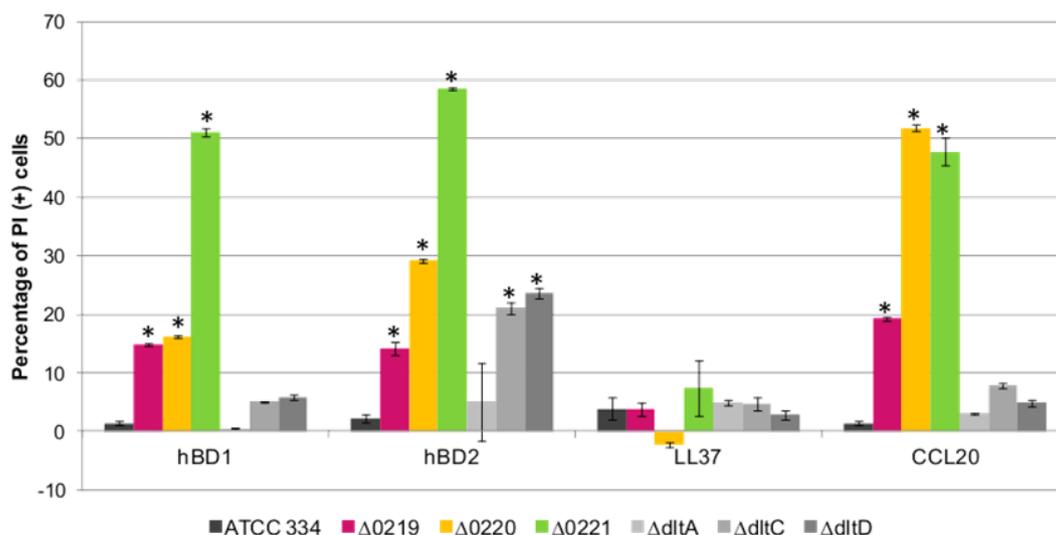
The resistance of the three mutants to environmental changes was assessed by the following procedure. Each of the three mutants was inoculated with ATCC 334 (2% of mutants and 98% of ATCC 334) during 16 h at 37°C in the following stressful media which replicate some of the ileal condition: (i) low-nutrient, acidic (pH 3.5), (ii) low-nutrient, bile supplemented (3 g.l<sup>-1</sup>), (iii) low-

nutrient, lysozyme-supplemented (10 g.l<sup>-1</sup>), then their viability was assessed by determining the ratio of mutant to the ATCC 334 and comparing it to the initial ratio (Fig. 9). The conditions demonstrated that the three mutants were not equal regarding their stress resistance. Δ0219 is altered for all the conditions whereas Δ0220 and Δ0221 are as resistant as the ATCC 334 to all conditions. These results suggest a specific involvement of the RR0219-HK0220 TCS in response to stress, independently modifications of cell surface architecture. These results also demonstrated that, taken separately, ileal physicochemical conditions are not enough to alter the three mutants.

Propidium iodide was used to assess the effect of antimicrobial peptides, hBD1, hBD2, LL37 and CCL20 on three mutants and *L. casei* ATCC 334 (Fig. 10). Lipoteichoic acid D-Alanylation reduces the negative net charge of the cell wall surface, allowing the resistance against antimicrobial cationic peptides (17). Thus, the *dltA*, *dltC* and *dltD* mutants were added to this test as predicted susceptible mutants. The Δ0219, Δ0220 and Δ0221 mutants for HBD1 and CCL20 are more susceptible than *dlt* mutants. The three mutants of operon are at least eight-fold more susceptible to hBD1 than the ATCC 334, and at least six-fold more susceptible to hBD2 and CCL20. However, the three mutants are less or equally susceptible than the ATCC 334 to LL37. The antimicrobial cationic peptides susceptibility could be responsible for the mutants decrease in the ligated rabbit ileal loop (4).



**Fig. 9:** Competition in stressful conditions between *L. casei* mutants (Δ0219, Δ0220 and Δ0221) and ATCC 334. The competition were carried out in diluted MRS broth (1/20) in 16 h at 37°C; Acid pH 3.5; Bile 3 g.l<sup>-1</sup>; Lysozyme 10 g.l<sup>-1</sup>. The percentage was determined by counting on Petri dishes. Three biological repeats were performed, and bars indicate standard deviations. Statistical analysis was performed using the unpaired Student *t* test: \*, *P* < 0.01; \*\*, *P* < 0.001.



**Fig. 10:** Percentage of PI (+) bacteria (permeabilized cell) after treatment with hBD1, hBD2, LL37 and CCL20. Values were obtained by flow cytometry. \* Indicate significant differences ( $P < 0.05$ )

## Discussion

In this study, we have presented how the RR0219-HK0220 TCS (two-component system encoded by LSEI\_0219-0220 gene) and the PBP0221 (penicillin-binding protein encoded by LSEI\_0221 gene) can operate in modeling the surface architecture and therefore modify bacterial ability to establish in the gut. The LSEI\_0219-0220 genes (encoding the TCS) are in operonic structure with LSEI\_0221 gene (encoding PBP0221), thus we hypothesized that it is related to cell wall biogenesis regulation. This was confirmed by the comparison of transcriptomic profiles of the Δ0219 and Δ0221 mutants showing a differential expression for 17 over the 50 cell wall-related genes tested. Also, other cell wall-related genes were differently regulated in the absence of the PBP, in a RR0219-HK0220 TCS-independent manner. As the RR2808-HK2807 TCS (encoded by LSEI\_2808-2807 genes) is structurally close to the WalK-WalR TCS, a TCS playing a role in cell wall metabolism (18-20), it is the best candidate to control these genes. In silico analysis of proteins structures lead to the assumption that RR0219-HK0220 TCS-dependent genes are rather dedicated to PG strand biogenesis whereas the other ones are rather dedicated to PG hydrolysis/degradation. On the opposite, LSEI\_0539 and LSEI\_0393 are overexpressed in Δ0221 mutant in stationary phase for both genes and in exponential phase for the second gene, thus they are likely regulated by the RR0219-HK0220 TCS. LSEI\_0539 encodes a cell-wall associated hydrolase with a NLPC-P60 and a glucosamidase domain which could be predicted to act as a N-acetylglucosamidase or an endopeptidase (21). LSEI\_0393 encodes a 1,4-beta-N-acetylmuramidase with a glyco\_hydro\_25 and a LysM domain which could act as a N-acetylmuramidase (21). Also, in addition to its cell wall related function, the RR0219-HK0220 TCS is required to resist to several stresses since the Δ0219 mutant was less able to survive to bile, acid or lysozyme than the ATCC 334 and the Δ0221 mutant. This particular phenotype is consistent with previous reports on *L. casei* (22).

The PG composition analysis made to confirm the PBP0221 predicted function did not lead to the expected structure. The PBP0221, annotated as a D-Alanyl-D-Alanine carboxypeptidase, acts as a L,D carboxypeptidase. The same difference between predictive and empiric functions was observed for DacB of *Lactococcus lactis* (23). Also, the differential regulation of some cell wall related genes in the TCS mutants did not trigger significant changes in the PG composition compared to the  $\Delta$ 0221 mutant. Also, the lack of PBP0221 did not alter cell morphology contrary to what we observed for other PG hydrolases (for instance in LSEI\_0281 mutant corresponding to p75 or *msp1* mutant in *L. casei* BL23 and in *L. rhamnosus* GG; (24, 25)), nor PG thickness. PG hydrolases group the whole PG degradation functions (N-acetyl muramidase, N-acetylglucosamidase, N-acetylmuramoyl-L-alanine amidase, endopeptidase and carboxypeptidase) required for the insertion of neosynthesized PG strands during cell growth and also for cellular division. According to their cleavage specificity, they can be autolysins, defined as enzymes that can, alone, induce bacterial lysis (26). Carboxypeptidases, therefore PBP0221, are not autolysin.

The global down-regulation of predicted PG hydrolases was observed in the  $\Delta$ 0221 mutant, particularly during stationary phase of growth. This lower need of PG hydrolases for mutants could be due to the PG structure itself (less cross linkings to hydrolyse) or to a lower PG turnover. The genes whose down-regulations are the strongest for the three mutants are LSEI\_0020, LSEI\_2029 and LSEI\_2226, which encode peptidoglycan hydrolase for which no precise function could be predicted. Their lower expression has no detectable incidence on the resistance to stress since their down regulation is identical for the three mutants whereas the TCS mutants ( $\Delta$ 0219 and  $\Delta$ 0220 mutants) are more susceptible to stress than the PBP mutant ( $\Delta$ 0221 mutant). Nevertheless, a differential post-transcriptional regulation between mutants cannot be excluded. Indeed, PG hydrolases activity can be modulate by their own O-glycosylation (27, 28), or by PG modifications such as NAG O-acetylation (29), NAG N-acetylation (30), or by teichoic acids d-alanylation (31, 32).

We assumed that the change in PG structure would also affect the barrier function of the cell wall and also modify the nature of the released muropeptides, thus the immune response. Above all, for the mutants, the two surface antigens (LSEI\_0020 and LSEI\_2029) are less present at the surface. This can be explained by a decrease of gene expression and/or by an alteration of the anchoring at the PG layer. The first proposal is validated for both genes. The protein encoded by LSEI\_0020 is the orthologous of p40 (or *msp2*) that can bind mucin (33). Nevertheless, in our conditions, no differences between mutants and ATCC 334 were observed for adhesion to mucin. In *L. rhamnosus*, p40 was attributed several intestinal epithelium homeostasis promoting functions. It can activate Akt (PKB; protein kinase B), inhibit cytokine-induced epithelial cell apoptosis, induce colon epithelial cell growth, decrease epithelial barrier damages induced by hydrogen peroxyde (34, 35). The protein encoded by LSEI\_2029 shares 72% identity with LGG\_02225 of *L. rhamnosus*, a protein that seems to be glycosylated (28). The mutant for LSEI\_2029 is not altered in gut establishment in the rabbit ligated ileal loop model (4), suggesting that the decrease of LSEI\_2029 alone is not enough to explain TCS and PBP mutants phenotypes in the rabbit gut. The phenotype of the three mutant can be explain either by the change in PG composition or the decrease of the two

antigens presentation or the combination of these two changes of the *L. casei* cell surface architecture.

The  $\Delta 0221$  mutant was not susceptible to stress in our conditions. *In vivo*, bacteria have to cope with various stressful environments but also with defense molecules produced by the host. We have tested the three mutants for their susceptibility to four antimicrobial peptides (AMP) commonly used to study interactions with bacteria. Human beta defensins are cationic AMP. hBD1 is constitutively expressed by epitheliums whereas hBD2 has to be activated, as it is especially done during inflammation (36). *L. casei* ATCC 334 is resistant to hBD1, hBD2, LL37 and CCL20 whereas *L. rhamnosus GG* and *L. delbrueckii* subsp. *bulgaricus* are susceptible to HBD2 (37, 38). All the tested mutants were resistant to LL37 at the tested concentrations. The *dlt* mutants (*dltA*, *dltC* and *dltD* mutants) altered in teichoid acid D-alanylation, used as susceptible controls, are mainly susceptible to hBD2, as it is the case for *L. rhamnosus dlt* mutant (15). The TCS-PBP mutants are more susceptible than the ATCC 334 to hBD1, hBD2 and CCL20 and also more susceptible than *dlt* mutant to hBD1 and CCL20. These results indicate that the increase susceptibility is due to the PBP absence and not related to the TCS. It makes sense since in Gram(+), the AMPs have first to pass through the PG. Inhibition of PG synthesis is an important mechanism of action of AMPs. PG precursors are activated and transported through the plasmic membrane, they can disturb PG synthesis but also its translocation and organization (transglycosylation et transpeptidation) (39).

In conclusion, we reported that the LSEI\_0219-0221 operon is composed of a functional TCS involved in the regulation of peptidoglycan synthesis and a PGH re-annotated by a L,D-carboxypeptidase instead of a D-Alanyl-D-Alanine carboxypeptidase. This operon is necessary for antigens presentation and a suitable architecture of the bacterial surface to establish in the gut and resist to host defenses.

## **Materials and methods**

### *Bacterial strains, plasmids and growth conditions*

Bacterial strains and plasmids used in this study are listed in table 2. *L. casei* ATCC 334 and mutants were grown statically at 37°C in MRS medium (Difco), supplemented with 5  $\mu\text{g}\cdot\text{ml}^{-1}$  erythromycin for mutants. Mutants  $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$ ,  $\Delta dltA$ ,  $\Delta dltC$ , and  $\Delta dltD$ , were obtained by transposon mutagenesis with  $P_{\text{junc}}\text{-TpaseI}_{1223}$  and identified by individual sequencing as previously described (4, 40, 41). *Escherichia coli* strains TG1 and BL21(DE3) were used as cloning and expression hosts, respectively. They were grown in LB medium at 37°C with shaking. Recombinant plasmids in *E. coli* were selected in LB medium containing 50  $\mu\text{g}\cdot\text{ml}^{-1}$  kanamycine.

### *DNA techniques*

All DNA manipulations were performed according to standard procedure (42). Plasmids were isolated by using Nucleospin plasmid miniprep kit (BioBasics). Ligation and restriction analysis were carried out according to the manufacturer's instructions. PCR was performed using 0.02 unit

of Phusion High-Fidelity DNA polymerase (Thermo Scientific), according to the manufacturer's recommendations, in an automatic thermocycler (Bio-Rad).

#### *RNA extractions and qRT-PCR-analysis*

Whole RNA extraction from 50 ml of culture in exponential phase ( $OD_{600nm} = 0.6$ ) and 10 ml in stationary phase ( $OD_{600nm} = 3.5$ ) was extracted after bead beating disruption using Tri reagent method (Sigma) and cDNA were synthesized as previously described (43). Quantitative reverse transcriptase PCRs (qRT-PCR) were performed in a CFX384 real-time detection system (Bio-Rad) using. The total volume of the PCR mixture was 15  $\mu$ l containing 1X SsoAdvanced<sup>TM</sup> Universal SYBR<sup>®</sup> Green Supermix (Bio-rad), 25 ng cDNA and 0.8  $\mu$ M of each primer (table 3). PCR amplification was one step at 95°C for 30 s followed by 40 cycles of 95°C for 5 s and 60°C for 10 s. An additional cooling step from 90 to 60°C (3°C/min) was performed to establish a melting curve in order to verify the homogeneity of the amplicon. All sample and primer combinations were assessed in three biological replicates with two technical replicates per biological replicate. In each run, a negative control using sterile water instead of cDNA was included. Normalized expression levels were calculated using the comparative critical threshold ( $\Delta\Delta Cq$ ) method in (44). The Real-Time data were processed with CFX Manager Software (Bio-Rad).

#### *Plasmid constructs and generation of E. coli strains overproducing RR0219 and cHK0220*

Full-length LSEI\_0219 was amplified using primers LSEI\_0219FbisNcoI (5'-ATACCATGGGCAAATTTTAATTGTTGA-3') and LSEI\_0219RXhoI (5'-GTGCTCGAGGGCCTCAACCTTATAGCCGA-3'), and the cytosolic domain of HK0220 (residues 137 to 396) was amplified using primers LSEI\_0220FHAMPNcoI (5'-ATACCATGGGCACCGTCAATTCAATGGCC-3') and LSEI\_0220RXhoI (5'-GTGCTCGAGTTTTGACTTTCCTGCTTCCT-3') from the *L. casei* genomic DNA. The forward primers contained the NcoI site in frame with the translation start codon and the reverse primers contained the XhoI site. After restriction digestion, the amplicons were ligated to NcoI-XhoI digested pET28a+ expression vector using T4 DNA ligase HC (Fermentas). The resulting ligation products were directly transformed by electroporation into the *E. coli* TG1 strain as described by Dower *et al.* 1988, to obtain pET0219 and pET0220HAMP constructs. Plasmids were isolated from selected transformants and subjected to sequencing reactions (GATC Biotech). The pET0219 and the pET0220HAMP were transformed into competent *E. coli* BL21 (DE3) cells by CaCl<sub>2</sub> procedure (45) and protein expression was induced at  $OD_{600nm} \approx 0.5$  in 1 L of LB with the addition of 0.5 mM IPTG for 3 h at 37°C.

#### *Cell extracts and protein purification*

*E. coli* or *L. casei* cells were washed and resuspended in 6 ml of 50 mM Tris-HCl, 1 mM PMSF, 150 mM NaCl, pH 7.5 and disrupted with 50  $\mu$ m glass beads in FastPrep system (QBiogene). The lysate was clarified by centrifugation and the supernatant was filtered through a 0.2  $\mu$ m filter (Millipore). For protein purification, the filtrates were applied to a Protino<sup>®</sup> Ni-NTA column (Macherey-Nagel). The RR0219 and cHK0220 proteins were eluted with a 20-250 imidazole

gradient (36 ml). Electrophoresis in 12.5% SDS-PAGE was performed to monitor the presence of the protein. Fractions containing purified protein were stored at 4°C in elution buffer. All proteins extracts were dialyzed against 4 L of buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl.

#### *Autophosphorylation and phosphotransfert assay*

Immediately before use, 0.5 ml of purified RR0219 and cHK0220 proteins (0.5 ml) were applied to a ultrafree-0.5 centrifugal filter unit (Millipore) and centrifuged for 10 min at 11000 rpm. The concentrated proteins were washed with 0.5 ml with 50 mM Tris-HCl, 250 mM NaCl, pH 8 using the filter unit. This washing was repeated two times. Protein concentrations were determined using Bradford protein assay kit (Bio Basic Inc.). For the autophosphorylation reaction, 6 µg of cHK0220 protein was incubated in a final volume of 20 µl phosphorylation buffer containing 50 mM Tris-HCl pH 7.5, 2.5 mM MnCl<sub>2</sub>, 5 mM MgCl<sub>2</sub>, 50 mM KCl, 20 µCi of [γ-<sup>32</sup>P]-ATP (3000 Ci/mmol) for 1, 5, 15, 30 and 60 min at 37°C. Reaction was started by addition of radiolabelled ATP and stopped by addition of 4 µl SDS loading buffer. For the phosphotransfert reaction, 10 µg of RR0219 protein was added after 60 min of cHK0220 protein autophosphorylation and samples were taken at 0, 15 and 60 min at 37°C and quenched by addition of 5 µl SDS loading buffer. The reaction mixtures were analyzed by a 12.5% SDS-PAGE, followed by autoradiography.

#### *Phosphorylation of RR0219 protein by small molecule phosphoryl donors*

[<sup>32</sup>P]-acetyl phosphate was prepared by incubating acetate kinase (0.4 units; Sigma A7437) in 55 µl of 25 mM Tris-HCl pH 7.5, 60 mM potassium acetate, 10 mM MgCl<sub>2</sub>, 100 µCi [γ-<sup>32</sup>P]-ATP (3000 Ci/mmol) for 30 min at 25°C. The [<sup>32</sup>P]-acetyl phosphate solution was mixed with 40 µl of 50 mM Tris-HCl pH 7.5, 2.5 mM MnCl<sub>2</sub>, 5 mM MgCl<sub>2</sub>, 50 mM KCl containing 40 µg of RR0219 protein. After incubation at 37°C, samples of 18 µl were taken at 15, 30, 60 and 120 min and quenched by addition of 2 µl of SDS loading buffer. The reaction mixtures were analyzed by a 12.5% SDS-PAGE, followed by autoradiography. RR0219 protein (3 µg) was incubated with 60 mM of acetyl phosphate 50 mM Tris-HCl pH 7.5, 2.5 mM MnCl<sub>2</sub>, 5 mM MgCl<sub>2</sub>, and 50 mM KCl, in a final volume of 10 µl for 5, 15, 30, 60 and 120 min at 37°C.

#### *EMSA*

EMSAs were carried out on cytosolic protein extract as previously described with some modifications (46). The 443-pb DNA probe, named Prom0219, corresponding to the whole LSEI\_0218-LSEI\_0219 intergenic region was PCR amplified with primers P0219F-189 (5'-GCTAAGGCTTTTGTCAATATGG-3') and P0219R254 (5'-CTTAGTAATTCGACGATTTC-3'). The fragment was 3' end-labeled with digoxigenin (DIG)-11-ddUTP with dig gel shift kit, 2<sup>nd</sup> generation (Roche) according to the manufacturer's recommendation. Before binding reactions, protein extract was heated for 10 min at 65°C, then it was incubated with 1 µl of 1/500 Prom0219 probe in 15 mM Tris-HCl pH 7.5, 5% glycerol (v/v), 2 mM EDTA, 50 mM NaCl, 25 mM KCl, 2.5 mM MgCl<sub>2</sub>, 2 mM DTT, 2.5 µg.ml<sup>-1</sup> of BSA (bovine serum albumin) and 2.5 µg.ml<sup>-1</sup> of salmon sperm DNA as an unspecific DNA competitor in a 15 µl reaction volume for 20 min at 30°C. The samples were resolved onto 5% (w/v) PAGE gel in 20 mM Tris-HCl, pH 7.8, 400 mM glycine, 1

mM EDTA at 4°C. The DNA was transferred to nylon membrane by electroblotting, probed with anti-DIG-AP antibody (Roche) and visualized using chemiluminescence.

#### *Transmission electron microscopy (TEM)*

Cells were pelleted by centrifugation and fixed overnight at 4°C in 3% glutaraldehyde (v/v), 2% formaldehyde (v/v), 5 mM MgCl<sub>2</sub>, 0.1 M phosphate buffer, pH 7.2. Then, cells were pelleted and embedded in 2.25% (w/v) low-gelling temperature agarose. Samples were fixed in 0.5% OsO<sub>4</sub> (m/v) in 0.1 M phosphate buffer pH 7.2 for 1 h at 4°C then dehydrated through graded concentration of ethanol then propylene oxide and embedded in Epon 812 (Merck, Lyon, France). Ultrathin sections were stained with PATAg method as described previously (47) and observed at 80 kV on a transmission electron microscope (H-7500, Hitachi, Tokyo).

#### *Peptidoglycan (PG) extraction and analysis*

*L. casei* ATCC 334 and mutants were purified and analyzed as described (48) with some modifications (49).

#### *Minimal inhibitory concentration (MIC) determination*

The MICs of penicillin, bacitracin, nisin and vancomycin against *L. casei* ATCC 334 and mutants were determined by a microdilution technique. Overnight cells cultures were diluted 1:100 in fresh MRS media and added in individual wells containing various concentrations of antibiotic (series of two-fold dilutions). The MIC was defined as the lowest antimicrobial concentration preventing visible turbidity after 24 h of incubation at 37°C.

#### *Microbial adhesion to solvents (MATS) test*

A MATS test was performed as described previously (50). Overnight cultures of each strain were washed and resuspended at OD<sub>600 nm</sub> = 0.4 in phosphate-buffered saline (PBS, pH 7.4) (A<sub>0</sub>). This suspension (3 ml) was mixed with 1 ml of solvent (chloroform, ethyl acetate or hexadecane) for 3 min with vortex. After 15 min (partition of phases), the DO<sub>600nm</sub> of the aqueous phase was measured (A<sub>1</sub>). The percentage of adhesion was calculated as follows: % adhesion = 100 x ((A<sub>0</sub>-A<sub>1</sub>) / A<sub>0</sub>). All experiments were repeated at least three times on three biological repeats.

#### *Triton X-100-induced autolysis*

*L. casei* strains were grown on MRS medium to mid-exponential phase, (OD<sub>600 nm</sub> = 0.8). Cells were harvested by centrifugation, washed, and then resuspended to OD<sub>600 nm</sub> of 1.0 in a 50 mM potassium phosphate buffer, pH 7.0 supplemented with 0.05% Triton X-100 (v/v). Cells suspensions were then transferred in 96-well sterile clear microplates (Nunc) and incubated 24 h at 30°C under agitation. Autolysis was monitored by measuring the OD<sub>600nm</sub>. All experiments were repeated at least three times on three biological repeats.

#### *Surface proteins extraction and identification*

Surface proteins were extracted by 5 M of LiCl according to the Smit *et al.* (2001) procedure (51, 52). Briefly, 20 ml of bacterial cultures were washed twice with ice distilled water. The cell pellets were resuspended in 2 ml of 5 M of LiCl and incubated 1 h at 37°C with shaking. Suspensions were centrifuged 10 000g at 10°C for 20 min then supernatants were precipitated with 4 ml of cold acetone and incubated 2 h at -20°C. The precipitated proteins were harvested by centrifugation (10 000g, 10 min at 4°C). The precipitates were resuspended in loading SDS buffer then boiled 3 min and analyzed by 12.5% SDS-PAGE. Briefly, for proteins identifications, proteins were reduced, alkylated then digested by trypsin (Promega) 3 h at 37°C. Tryptic digests were separated and analyzed by liquid chromatography coupled to tandem mass spectroscopy (LC-MS/MS) (nanoLC-ESI-Trap, Orbitrap ELITE, ThermoScientific). Protein identification was performed using Proteome Discoverer 1.4 software against a protein database of *L. casei* ATCC 334 strain (GenBank: NC\_008526.1 and NC\_008502.1).

#### *Challenge in stressful conditions*

Overnight cultures were assembled to get a proportion of 2% of mutant and 98% of ATCC 334. The resulting mixed cultures were washed with PBS pH 7.4, resuspended to a final OD<sub>600 nm</sub> of 0.1 and challenged during 16 h at 37°C in stressful conditions in diluted MRS (1/20): acidity (pH 3.5), bile supplemented (3g.l<sup>-1</sup>), lysozyme supplemented (10g.l<sup>-1</sup>). The concentrations of mutant relative to ATCC 334 were determined by counting (CFU.ml<sup>-1</sup>) on MRS + Ery 5 µg.ml<sup>-1</sup> and MRS plates.

#### *Susceptibility to cationic peptides*

Cationic peptides susceptibility tests were performed as previously described (38). 50 µl of bacterial cultures diluted in 50 mM *N*-2-hydroxyethylpiperazine-*N*9-2-ethanesulfonic acid (Hepes) buffer pH 7.5 at 4.10<sup>7</sup> CFU.ml<sup>-1</sup> were incubated with 50 µl of hBD1, hBD2, LL37 or CCL20 (abcam) at 20 µg.ml<sup>-1</sup> in Hepes buffer for 1 h at 37°C with 5% CO<sub>2</sub>. Then, treated and non-treated bacteria were incubated 5 min at room temperature with 400 µl of propidium iodide (PI) at 5 µg.ml<sup>-1</sup>. The ratio of PI(+) cells was assessed by flow cytometry (FACSCanto, Diva software; Becton Dickinson, Mountain View, CA, USA). For each sample, 10 000 events were analysed.

#### **Acknowledgements**

This study was supported by the European Research Council Advanced grant HOMEOPITH to PJS, Grant agreement no: 232798. PJS is a Howard Hughes Medical Institute Foreign Scholar. HS was supported by a grant from the Ministère de l'Enseignement Supérieur et de la Recherche. We thank the Pasteur Institute Microorganism Collection for the *L. casei* Type strain (CIP 107868, ATCC 334). We wish to express particular thanks to Clément Petitgonnet for its technical contribution, Geraldine Lucchi and Patrick Ducoroy (CLIPP Clinical Innovation Proteomic Platform, IFR 100 Santé-STIC, Université de Bourgogne, Dijon, France) for carrying out protein identifications, and Jeannine Lherminier and Joël Michel (Centre de Microscopie INRA/uB, DimaCell Plateform, Dijon, France) for their advice and help regarding TEM.

Table 2: Bacterial strains and plasmids

Strains and plasmids	Genotype and/or relevant feature(s)	Source or reference
<b>Strains</b>		
<i>L. casei</i> ATCC 334	Wild type, CIP 107868, genome sequenced	Collection Institut Pasteur, France
Δ0219	ATCC 334 LSEI_0219::pVII110, Erm <sup>R</sup>	(4)
Δ0220	ATCC 334 LSEI_0220::pVII110, Erm <sup>R</sup>	(4)
Δ0221	ATCC 334 LSEI_0221::pVII110, Erm <sup>R</sup>	(4)
ΔdltA	ATCC 334 LSEI_0794::pVII110, Erm <sup>R</sup>	(4)
ΔdltC	ATCC 334 LSEI_0796::pVII110, Erm <sup>R</sup>	(4)
ΔdltD	ATCC 334 LSEI_0797::pVII110, Erm <sup>R</sup>	(4)
<i>E. coli</i> TG1	<i>SupE hsdΔ5thi Δ(lac-proAB)F' [tra D36 pro AB<sup>+</sup> lac I<sup>f</sup> lacZΔM15]</i>	(53)
<i>E. coli</i> BL21 (DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (rB<sup>-</sup>mB<sup>-</sup>) gal dcm</i> (DE3)	Invitrogen
<b>Plasmids</b>		
pET28a+	Kan <sup>r</sup> , vector for overexpression of His-tagged proteins using the T7 bacteriophage promoter	Novagen
pET0219bis	pET28a+ containing LSEI_0219 gene between NcoI and XhoI sites to overproduce RR0219 His <sub>6</sub> tag	This work
pET0220HAMP	pET28a+ containing truncated LSEI_0220 gene coding for residues 137-396 of the HK protein between NcoI and XhoI sites to overproduce cHK0220 His <sub>6</sub> tag	This work

Table 3: qRT-PCR primers

Name Forward	Sequence (5' → 3')	Name Reverse	Sequence (5' → 3')
LSEI 0020 F	GTGATTCTGGGTTTCCAATG	LSEI 0020 R	TTGATGCCGTGACAGTTGCC
LSEI 0143 F	ATCACCAAATCGCCGTAGTC	LSEI 0143 R	GTCATTGAAGCCAACAACCTG
LSEI 0219 F	CGGTAAGGAAGCCATTACTAA	LSEI 0219 R	CAACATCCATCCAGACATATTA
LSEI 0220 F	TCGCAGTGAAGAAGAGCTACTC	LSEI 0220 R	TGGTCGATTTAACCTTGGTGTA
LSEI 0221 F	CAGGTGGCACAACATTTACTG	LSEI 0221 R	CGCATCACTGGTTCCCTTAC
LSEI 0223 F	CGGCTGCTATTATTGGTGCT	LSEI 0223 R	CGGCAACTTTCGTATGCTCC
LSEI 0233 F	GACTGGTGGATTGCTTTGGT	LSEI 0233 R	GGCTACCCGTCACATTACC
LSEI 0234 F2	CTGGAAGTGGGCAGCAGTTA	LSEI 0234 R2	TGTTAGCGGAATCTGGCGAA
LSEI 0236 F	CTCAAAGAAGCACGGACGAG	LSEI 0236 R	ATGACGACTGAATGCCAAGC
LSEI 0249 F	CCGTAACCCAAACCAGACC	LSEI 0249 R	CGACAGAAACAACCACCAAC
LSEI 0251 F	CCAAAGGCGATTTAGTTGCTT	LSEI 0251 R	CAATGAGATCCAGCCACG
LSEI 0281 F	GCCTGGTGATTGCTCTTCT	LSEI 0281 R	CTGGTTGTGGTGCCTGAATC
LSEI 0314 F	GCCGTCAGTGTCTTGCTTT	LSEI 0314 R	GGAACCGCCACAATCTC
LSEI 0393 F	GTAATGGCAGCAGATGTGGA	LSEI 0393 R	GCTTTGACCGTTTGGAGAAA
LSEI 0435 F	GACGCCTCTATGTCGAGCTT	LSEI 0435 R	GCCGTTTACCCTTCAAACAA
LSEI 0539 F	AATTGGAGCAGATGCCAGT	LSEI 0539 R	CTTCGCTTCGGTAATGATGG
LSEI 0599 F	GCTACATCTGGTATGGTGTGG	LSEI 0599 R	ATGCTTCCCTTTGGTGTCTG
LSEI 0708 F	GGTCTGTTTCATTTCATCTGCT	LSEI 0708 R	TCGCTTACTCAAGGCTGTCA
LSEI 0799 F2	GGAAATGACCGCTGAAAGAAA	LSEI 0799 R2	CATCCGTTTGCTTCCAAC
LSEI 0810 F	AAGTGC GGCTTGTATTGGAA	LSEI 0810 R	CGGTTTATTGGCGGCTAT
LSEI 0906 F	GCGACGAGGAAAGAAACAAT	LSEI 0906 R	AAATCGGCAAATACCTGTGG
LSEI 1003 F	CGCCACCTTTACGCTCAA	LSEI 1003 R	AAGACCGACCCGCAAGAAG
LSEI 1072 F	CCAAGTGTGCTGATGGGTA	LSEI 1072 R	CTGCTCGTTGTCGGCTTCT
LSEI 1089 F	GGCGATCCATCAACTAGGC	LSEI 1089 R	TAACACAAATAGACTCAGGGC
LSEI 1110 F	CCTTACCAAACGAGCGAGAA	LSEI 1110 R	TTCAAGCGACGATACCCTTC
LSEI 1116 F	GTATGACAACGACAGTGAATCG	LSEI 1116 R	TCCTTAGCGGCACGAACT
LSEI 1152 F3	TTCCCACACACAACCACTTT	LSEI 1152 R3	GTCATCAACCGCTGTTTCAG
LSEI 1153 F	GCGACTGTAAGGCGTAAACC	LSEI 1153 R	TGGATTTGGGACGGCAAT
LSEI 1269 F	CGTGACTCGCTTGTGAAGA	LSEI 1269 R	GGGTGATACTTGCCACTTTGA
LSEI 1270 F	TGTTGGCAGTGGTTGCTATC	LSEI 1270 R	CTGGCTGTTTCGGTGACATAA
LSEI 1271 F	GATTATCACTGAACCCGAGTT	LSEI 1271 R	ATCAAGCCAATCAGCGTGGT
LSEI 1272 F	TGCCCAAAGTCTAACGAAACA	LSEI 1272 R	TCTCCATCAAGGTCACAACC
LSEI 1314 F	ACTTTCGGTCAAGGCAGATG	LSEI 1314 R	CCACGACCAATGTAATCAAGA
LSEI 1351 F	AAACGCATCCAGTCAATCGTG	LSEI 1351 R	CAGAGTCGCCAATAGCAGTG
LSEI 1481 F	GTGATGCCGAACCTCAAACA	LSEI 1481 R	GGTCCAAACCGAAATGACTC
LSEI 1536 F	TGACCAACCTTCGCAATG	LSEI 1536 R	TGACCGCATACCAATCATCT
LSEI 1554 F	GGCAAGAACCAGTCGGTATT	LSEI 1554 R	CGCAGCATCTCTGTTATGA
LSEI 1662 F	ACCTGCGTTTGACTCTTGG	LSEI 1662 R	TGGCTATGCTCGGAATGAC
LSEI 1711 F	CGTGGAAGTAAGCATCAAAGC	LSEI 1711 R	TTATGGGATTAGTGACCGTGA
LSEI 1725 F	GCTTCTCGTTGTTAGCGTCA	LSEI 1725 R	CGGGTCTCTGCTTGGTTTA
LSEI 1802 F	CTGTGTGCTTCAGGTTGAGC	LSEI 1802 R	CGTGTACTTGCTGGTGTGA
LSEI 1805 F	CGTCAGAAAGCGAAGCACTC	LSEI 1805 R	GGACAGGCAAGCAAGTTACC
LSEI 1909 F	GTAGGTCTTGCCGCTGTCA	LSEI 1909 R	TGCTGGTGGACTTGATGGTA
LSEI 2029 F	CGGTGCTGGTTGAAGAAGTAG	LSEI 2029 R	AAAGATAAGTTGGCTCAGTTGG
LSEI 2137 F	AAGTTATGGGTGTCGTCAG	LSEI 2137 R	GGGATTGTGCGTCAAGTTAC
LSEI 2226 F	TAGCAGCAACGCCACTGTAG	LSEI 2226 R	ACGACACAACCAACCCAAAG
LSEI 2550 F2	CTGGATGCCGATTGATTACA	LSEI 2550 R2	GTCTGGTGGGTGTTGACG
LSEI 2553 F	GTTGCCAGTTTGTGTTGTCG	LSEI 2553 R	TTTGTGGTTGATAGAGCGATAG
LSEI 2560 F	GCACGATAGGTCACACCGTA	LSEI 2560 R	GATTCCAACCACACCGTTTC
LSEI 2563 F2	CAATCCGTGACTCGCTTTC	LSEI 2563 R2	GCACCTATTATCCGCTTGATGA
LSEI 2570 F	GTGTTCAATTGCTGACGGAAG	LSEI 2570 R	TAGGGCGGTTTGGTAAAGC
LSEI 2677 F	CTAAGCGGGTTTGAATCGTAT	LSEI 2677 R	GATGATGAGTCTGGGCATCA
LSEI 2773 F	TGGGCATCAGGGACATCT	LSEI 2773 R	CATAGGCTCGTTGGTCAAGC
LCLEUSF	GCTGGATGCTGGTATTGCTT	LCLEUSR	ATCACGCAGTTTGCCTTCAT
LCRECGF	CGTAGCGTGATTCTGGTGAC	LCRECGR	TTCCAAAGATTGCTGCTTCA

Primers names correspond to names of target genes.

## References

1. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Diaz-Muniz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D. 2006. Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* **103**:15611-15616.
2. van de Guchte M, Serror P, Chervaux C, Smokvina T, Ehrlich SD, Maguin E. 2002. Stress responses in lactic acid bacteria. *Antonie Van Leeuwenhoek* **82**:187-216.
3. Chapot-Chartier M-P, Kulakauskas S. 2014. Cell wall structure and function in lactic acid bacteria. *Microbial Cell Factories* **13**:S9.
4. Licandro-Seraut H, Scornec H, Pedron T, Cavin JF, Sansonetti PJ. 2014. Functional genomics of *Lactobacillus casei* establishment in the gut. *Proc Natl Acad Sci U S A* **111**:E3101-3109.
5. Cai H, Thompson R, Budinich MF, Broadbent JR, Steele JL. 2009. Genome sequence and comparative genome analysis of *Lactobacillus casei*: insights into their niche-associated evolution. *Genome Biol Evol* **1**:239-257.
6. Handwerger S, Pucci MJ, Volk KJ, Liu J, Lee MS. 1994. Vancomycin-resistant *Leuconostoc mesenteroides* and *Lactobacillus casei* synthesize cytoplasmic peptidoglycan precursors that terminate in lactate. *J Bacteriol* **176**:260-264.
7. Klare I, Konstabel C, Werner G, Huys G, Vankerckhoven V, Kahlmeter G, Hildebrandt B, Muller-Bertling S, Witte W, Goossens H. 2007. Antimicrobial susceptibilities of *Lactobacillus*, *Pediococcus* and *Lactococcus* human isolates and cultures intended for probiotic or nutritional use. *J Antimicrob Chemother* **59**:900-912.
8. Mackey T, Lejeune V, Janssens M, Wauters G. 1993. Identification of vancomycin-resistant lactic bacteria isolated from humans. *J Clin Microbiol* **31**:2499-2501.
9. Balandino G, Milazzo I, Fazio D. 2008. Antibiotic susceptibility of bacterial isolates from probiotic products available in Italy. *Microbial Ecol Health Dis* **20**:199-203.
10. Wilks M, Wiggins R, Whiley A, Hennessy E, Warwick S, Porter H, Corfield A, Millar M. 2004. Identification and H<sub>2</sub>O<sub>2</sub> production of vaginal lactobacilli from pregnant women at high risk of preterm birth and relation with outcome. *Journal of Clinical Microbiology* **42**:713-717.
11. Rossetti L, Carminati D, Zago M, Giraffa G. 2009. A Qualified Presumption of Safety approach for the safety assessment of Grana Padano whey starters. *Int J Food Microbiol* **130**:70-73.
12. Glauner B. 1988. Separation and quantification of mucopeptides with high-performance liquid chromatography. *Anal Biochem* **172**:451-464.
13. Munoz-Provencio D, Llopis M, Antolin M, de Torres I, Guarner F, Perez-Martinez G, Monedero V. 2009. Adhesion properties of *Lactobacillus casei* strains to resected intestinal fragments and components of the extracellular matrix. *Arch Microbiol* **191**:153-161.
14. Raychaudhuri D, Chatterjee AN. 1985. Use of resistant mutants to study the interaction of triton X-100 with *Staphylococcus aureus*. *J Bacteriol* **164**:1337-1349.
15. Perea Velez M, Verhoeven TL, Draing C, Von Aulock S, Pfitzenmaier M, Geyer A, Lambrichts I, Grangette C, Pot B, Vanderleyden J, De Keersmaecker SC. 2007. Functional analysis of D-alanylation of lipoteichoic acid in the probiotic strain *Lactobacillus rhamnosus* GG. *Appl Environ Microbiol* **73**:3595-3604.
16. Yan F, Liu L, Dempsey PJ, Tsai YH, Raines EW, Wilson CL, Cao H, Cao Z, Polk DB. 2013. A *Lactobacillus rhamnosus* GG-derived soluble protein, p40, stimulates ligand release from intestinal epithelial cells to transactivate epidermal growth factor receptor. *J Biol Chem* **288**:30742-30751.
17. Neuhaus FC, Baddiley J. 2003. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**:686-723.
18. Dubrac S, Boneca IG, Poupel O, Msadek T. 2007. New insights into the Walk/WalR (YycG/YycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in *Staphylococcus aureus*. *J Bacteriol* **189**:8257-8269.
19. Bisicchia P, Noone D, Lioliou E, Howell A, Quigley S, Jensen T, Jarmer H, Devine KM. 2007. The essential YycFG two-component system controls cell wall metabolism in *Bacillus subtilis*. *Mol Microbiol* **65**:180-200.
20. Delaune A, Poupel O, Mallet A, Coic YM, Msadek T, Dubrac S. 2011. Peptidoglycan crosslinking relaxation plays an important role in *Staphylococcus aureus* WalkR-dependent cell viability. *PLoS One* **6**:e17054.

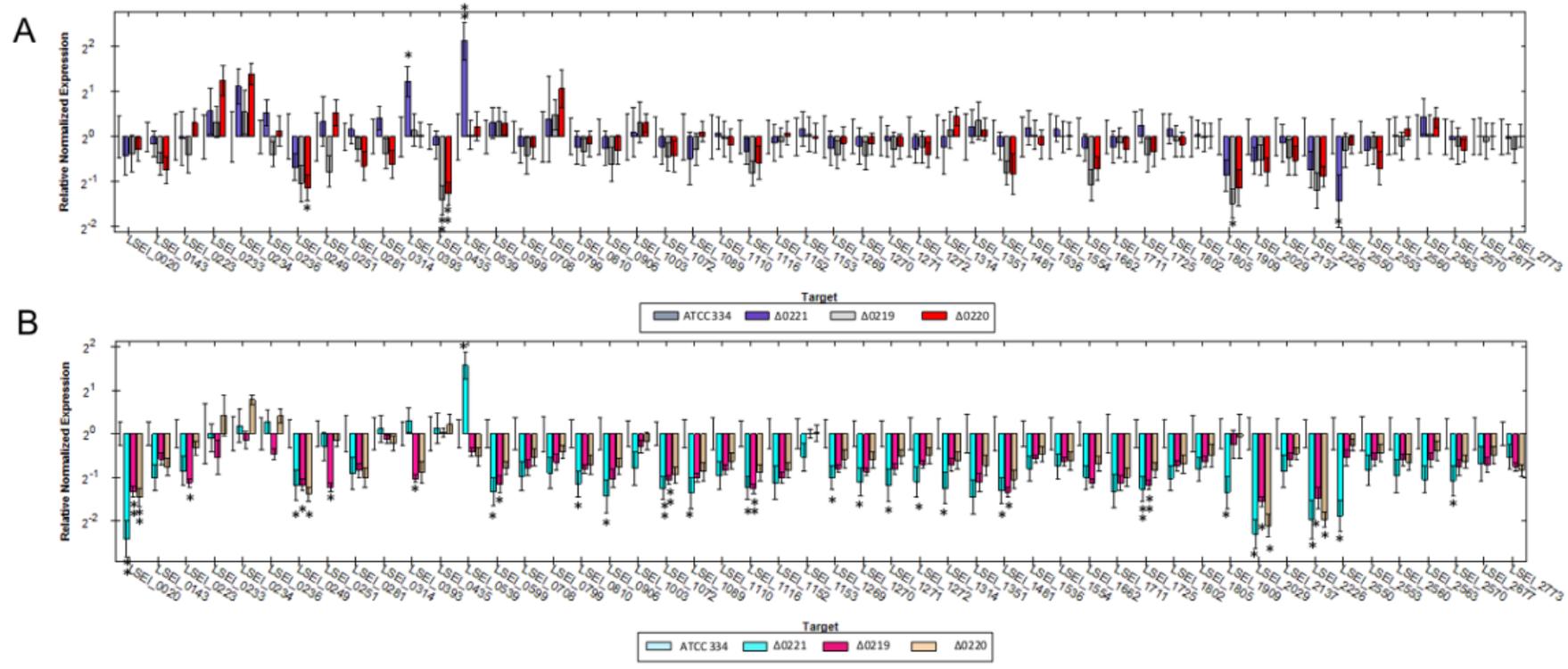
21. Layec S, Decaris B, Leblond-Bourget N. 2008. Diversity of Firmicutes peptidoglycan hydrolases and specificities of those involved in daughter cell separation. *Res Microbiol* **159**:507-515.
22. Alcantara C, Revilla-Guarinos A, Zuniga M. 2011. Influence of two-component signal transduction systems of *Lactobacillus casei* BL23 on tolerance to stress conditions. *Appl Environ Microbiol* **77**:1516-1519.
23. Courtin P, Miranda G, Guillot A, Wessner F, Mezange C, Domakova E, Kulakauskas S, Chapot-Chartier MP. 2006. Peptidoglycan structure analysis of *Lactococcus lactis* reveals the presence of an L,D-carboxypeptidase involved in peptidoglycan maturation. *J Bacteriol* **188**:5293-5298.
24. Claes IJ, Schoofs G, Regulski K, Courtin P, Chapot-Chartier MP, Rolain T, Hols P, von Ossowski I, Reunanen J, de Vos WM, Palva A, Vanderleyden J, De Keersmaecker SC, Lebeer S. 2012. Genetic and biochemical characterization of the cell wall hydrolase activity of the major secreted protein of *Lactobacillus rhamnosus* GG. *PLoS One* **7**:e31588.
25. Regulski K, Courtin P, Meyrand M, Claes IJ, Lebeer S, Vanderleyden J, Hols P, Guillot A, Chapot-Chartier MP. 2012. Analysis of the peptidoglycan hydrolase complement of *Lactobacillus casei* and characterization of the major gamma-D-glutamyl-L-lysyl-endopeptidase. *PLoS One* **7**:e32301.
26. Shockman GD, Høltje J-V. 1994. Microbial peptidoglycan (murein) hydrolases. *New Compr. Biochem.* **27**:131-166.
27. Rolain T, Bernard E, Beaussart A, Degand H, Courtin P, Egge-Jacobsen W, Bron PA, Morsomme P, Kleerebezem M, Chapot-Chartier MP, Dufrene YF, Hols P. 2013. O-glycosylation as a novel control mechanism of peptidoglycan hydrolase activity. *J Biol Chem* **288**:22233-22247.
28. Lebeer S, Claes IJ, Balog CI, Schoofs G, Verhoeven TL, Nys K, von Ossowski I, de Vos WM, Tytgat HL, Agostinis P, Palva A, Van Damme EJ, Deelder AM, De Keersmaecker SC, Wuhrer M, Vanderleyden J. 2012. The major secreted protein Msp1/p75 is O-glycosylated in *Lactobacillus rhamnosus* GG. *Microb Cell Fact* **11**:15.
29. Bernard E, Rolain T, Courtin P, Guillot A, Langella P, Hols P, Chapot-Chartier MP. 2011. Characterization of O-acetylation of N-acetylglucosamine: a novel structural variation of bacterial peptidoglycan. *J Biol Chem* **286**:23950-23958.
30. Meyrand M, Boughammoura A, Courtin P, Mezange C, Guillot A, Chapot-Chartier MP. 2007. Peptidoglycan N-acetylglucosamine deacetylation decreases autolysis in *Lactococcus lactis*. *Microbiology* **153**:3275-3285.
31. Steen A, Palumbo E, Deghorain M, Cocconcelli PS, Delcour J, Kuipers OP, Kok J, Buist G, Hols P. 2005. Autolysis of *Lactococcus lactis* is increased upon D-alanine depletion of peptidoglycan and lipoteichoic acids. *J Bacteriol* **187**:114-124.
32. Palumbo E, Deghorain M, Cocconcelli PS, Kleerebezem M, Geyer A, Hartung T, Morath S, Hols P. 2006. D-alanyl ester depletion of teichoic acids in *Lactobacillus plantarum* results in a major modification of lipoteichoic acid composition and cell wall perforations at the septum mediated by the Acm2 autolysin. *J Bacteriol* **188**:3709-3715.
33. Bauerl C, Perez-Martinez G, Yan F, Polk DB, Monedero V. 2010. Functional analysis of the p40 and p75 proteins from *Lactobacillus casei* BL23. *J Mol Microbiol Biotechnol* **19**:231-241.
34. Yan F, Cao H, Cover TL, Whitehead R, Washington MK, Polk DB. 2007. Soluble proteins produced by probiotic bacteria regulate intestinal epithelial cell survival and growth. *Gastroenterology* **132**:562-575.
35. Seth A, Yan F, Polk DB, Rao RK. 2008. Probiotics ameliorate the hydrogen peroxide-induced epithelial barrier disruption by a PKC- and MAP kinase-dependent mechanism. *Am J Physiol Gastrointest Liver Physiol* **294**:G1060-1069.
36. Braff MH, Bardan A, Nizet V, Gallo RL. 2005. Cutaneous defense mechanisms by antimicrobial peptides. *J Invest Dermatol* **125**:9-13.
37. De Keersmaecker SC, Braeken K, Verhoeven TL, Perea Velez M, Lebeer S, Vanderleyden J, Hols P. 2006. Flow cytometric testing of green fluorescent protein-tagged *Lactobacillus rhamnosus* GG for response to defensins. *Appl Environ Microbiol* **72**:4923-4930.
38. Hugo AA, Tymczynsyn EE, Gomez-Zavaglia A, Perez PF. 2012. Effect of human defensins on lactobacilli and liposomes. *J Appl Microbiol* **113**:1491-1497.
39. Yeaman MR, Yount NY. 2003. Mechanisms of antimicrobial peptide action and resistance. *Pharmacol Rev* **55**:27-55.
40. Scornec H, Tichit M, Bouchier C, Pedron T, Cavin JF, Sansonetti PJ, Licandro-Seraut H. 2014. Rapid 96-well plates DNA extraction and sequencing procedures to identify genome-wide transposon insertion sites in a difficult to lyse bacterium: *Lactobacillus casei*. *J Microbiol Methods*.

41. **Licandro-Seraut H, Brinster S, van de Guchte M, Scornec H, Maguin E, Sansonetti P, Cavin JF, Serror P.** 2012. Development of an efficient *in vivo* system ( $P_{\text{junc-TpaselS}_{1223}}$ ) for random transposon mutagenesis of *Lactobacillus casei*. *Appl Environ Microbiol* **78**:5417-5423.
42. **Sambrook J, Fritsch EF, Maniatis T.** 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor, NY.
43. **Licandro-Seraut H, Gury J, Tran NP, Barthelmebs L, Cavin JF.** 2008. Kinetics and intensity of the expression of genes involved in the stress response tightly induced by phenolic acids in *Lactobacillus plantarum*. *J Mol Microbiol Biotechnol* **14**:41-47.
44. **Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J.** 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol* **8**:R19.
45. **Cohen SN, Chang AC, Hsu L.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci U S A* **69**:2110-2114.
46. **Gury J, Barthelmebs L, Tran NP, Divies C, Cavin JF.** 2004. Cloning, deletion, and characterization of PadR, the transcriptional repressor of the phenolic acid decarboxylase-encoding *padA* gene of *Lactobacillus plantarum*. *Appl Environ Microbiol* **70**:2146-2153.
47. **Thiéry J-P.** 1967. Mise en évidence des polysaccharides sur coupes fines en microscopie électronique. *J. Microscopie* **6**:987-1018.
48. **de Jonge BL, Chang YS, Gage D, Tomasz A.** 1992. Peptidoglycan composition of a highly methicillin-resistant *Staphylococcus aureus* strain. The role of penicillin binding protein 2A. *J Biol Chem* **267**:11248-11254.
49. **Girardin SE, Boneca IG, Viala J, Chamaillard M, Labigne A, Thomas G, Philpott DJ, Sansonetti PJ.** 2003. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* **278**:8869-8872.
50. **Bellon-Fontaine M-N, Rault J, van Oss CJ.** 1996. Microbial adhesion to solvents: a novel method to determine the electron-donor/electron-acceptor or Lewis acid-base properties of microbial cells. *Colloids and Surf. B: Biointerfaces* **7**:47-53.
51. **Zhang YC, Zhang LW, Tuo YF, Guo CF, Yi HX, Li JY, Han X, Du M.** 2010. Inhibition of *Shigella sonnei* adherence to HT-29 cells by lactobacilli from Chinese fermented food and preliminary characterization of S-layer protein involvement. *Res Microbiol* **161**:667-672.
52. **Smit E, Oling F, Demel R, Martinez B, Pouwels PH.** 2001. The S-layer protein of *Lactobacillus acidophilus* ATCC 4356: identification and characterisation of domains responsible for S-protein assembly and cell wall binding. *J Mol Biol* **305**:245-257.
53. **Gibson TJ.** 1984. Studies on the Epstein-Barr virus genome. Ph.D. thesis Cambridge University, Cambridge, United Kingdom.

**Supplemental data:**

**Table S1:** Tested genes and corresponding annotations in *L. casei* genome

gene	function
LSEI_0020	surface antigen
LSEI_0143	D-alanine-D-alanine ligase and related ATP-grasp enzymes
LSEI_0223	UDP-N-acetylmuramyl tripeptide synthase
LSEI_0233	Glycosyltransferases involved in cell wall biogenesis
LSEI_0234	Glycosyltransferases involved in cell wall biogenesis
LSEI_0236	Lysozyme M1 (1,4-beta-N-acetylmuramidase)
LSEI_0249	Cell wall-associated hydrolase
LSEI_0251	Glycosyltransferases involved in cell wall biogenesis
LSEI_0281	Cell wall-associated hydrolases (invasion-associated proteins)
LSEI_0314	Cell division protein FtsI/penicillin-binding protein 2
LSEI_0393	Lysozyme M1 (1,4-beta-N-acetylmuramidase)
LSEI_0435	Beta-lactamase class C and other penicillin binding proteins
LSEI_0539	Cell wall-associated hydrolases (invasion-associated proteins)
LSEI_0599	1,4-beta-N-acetylmuramidase
LSEI_0708	Glycosyltransferases involved in cell wall biogenesis
LSEI_0799	Xanthosine triphosphate pyrophosphatase
LSEI_0810	Membrane carboxypeptidase (penicillin-binding protein)
LSEI_0906	Uncharacterized bacitracin resistance protein
LSEI_1003	UDP-N-acetylmuramate dehydrogenase
LSEI_1072	Lysozyme M1 (1,4-beta-N-acetylmuramidase)
LSEI_1089	Glycosyltransferases involved in cell wall biogenesis
LSEI_1110	Glycosyltransferases involved in cell wall biogenesis
LSEI_1116	Muramidase (flagellum-specific)
LSEI_1152	Beta-lactamase class C and other penicillin binding proteins
LSEI_1153	UDP-N-acetylmuramyl tripeptide synthase
LSEI_1269	Cell division protein FtsI/penicillin-binding protein 2
LSEI_1270	UDP-N-acetylmuramyl pentapeptide phosphotransferase/UDP-N-acetylglucosamine-1-phosphate transferase
LSEI_1271	UDP-N-acetylmuramoylalanine-D-glutamate ligase
LSEI_1272	UDP-N-acetylglucosamine:LPS N-acetylglucosamine transferase
LSEI_1314	Bacterial cell division membrane protein
LSEI_1351	Predicted acyltransferases
LSEI_1481	Membrane carboxypeptidase (penicillin-binding protein)
LSEI_1536	N-acetylmuramoyl-L-alanine amidase
LSEI_1554	N-formylglutamate amidohydrolase
LSEI_1662	Cell division protein FtsI/penicillin-binding protein 2
LSEI_1711	UDP-N-acetylmuramate-alanine ligase
LSEI_1725	Membrane carboxypeptidase (penicillin-binding protein)
LSEI_1802	Beta-lactamase class C and other penicillin binding proteins
LSEI_1805	Beta-lactamase class C and other penicillin binding proteins
LSEI_1909	1,4-beta-N-acetylmuramidase
LSEI_2029	Cell wall-associated hydrolases (invasion-associated proteins), surface antigen
LSEI_2137	Beta-lactamase class C and other penicillin binding proteins
LSEI_2226	Cell wall-associated hydrolase
LSEI_2550	Uncharacterized protein conserved in bacteria
LSEI_2553	hypothetical protein
LSEI_2560	Alanine racemase
LSEI_2563	UDP-N-acetylmuramyl pentapeptide synthase
LSEI_2570	UDP-N-acetylglucosamine enolpyruvyl transferase
LSEI_2677	Glycosyltransferases involved in cell wall biogenesis
LSEI_2773	Beta-lactamase class C and other penicillin binding proteins

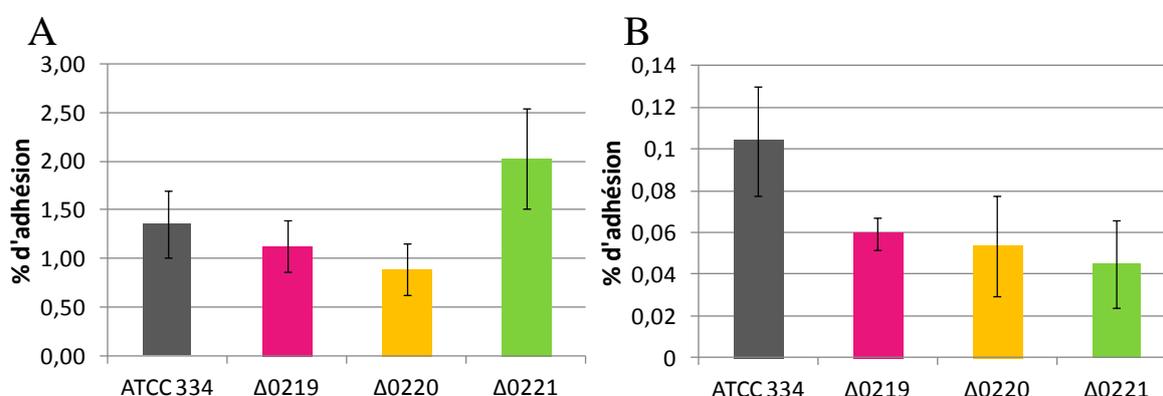


**Fig. S1** : Relative transcript levels of genes implicated in peptidoglycan synthesis in Δ0219, Δ0220 and Δ0221 mutants compared to the ATCC 334 (WT) grown either until exponential (A) or stationary phase (B). Values are the mean RTL obtained for three biological repeats. Statistical analysis was performed; \*\*, P<0.01; \*, P<0.05.

## Résultats complémentaires à l'article 4.

### Adhésion à une surface biotique et abiotique

Des tests d'adhésions des mutants  $\Delta 0219$ ,  $\Delta 0220$ ,  $\Delta 0221$  et de la souche parentale *L. casei* ATCC 334 ont été réalisés sur des plaques polystyrène 6 puits non traitées recouvertes ou non de mucine d'estomac de porc (Sigma). L'adhésion a été testée après une incubation de 3h30 à 37°C pour les plaques recouvertes de mucine et après 48 h à 37°C pour les plaques non recouvertes. Le pourcentage d'adhésion a été calculé en utilisant comme témoin les cultures initialement ensemencées et incubées dans les mêmes conditions que les plaques. Les deux types de tests d'adhésion réalisés ont montré un faible pourcentage d'adhésion de *L. casei* : environ 1% pour la mucine et 0,1% pour les plaques non recouvertes. Aucune différence significative d'adhésion entre les mutants et la souche parentale n'a été observée (Figure 34).



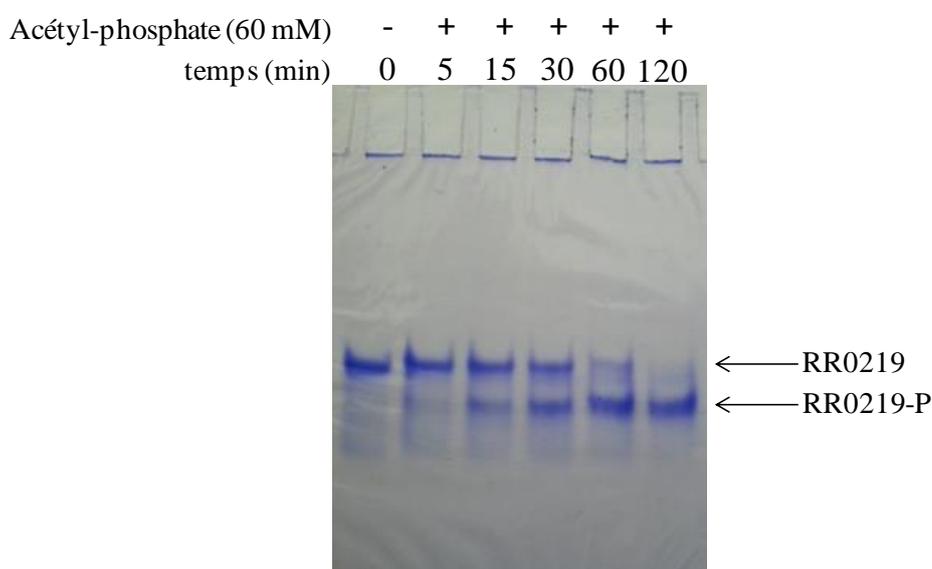
**Figure 34 : Adhésion des mutants de l'opéron ( $\Delta 0219$ ,  $\Delta 0220$  et  $\Delta 0221$ ) et de *L. casei* ATCC 334 sur plaques recouvertes ou non de mucine.**

Le % d'adhésion a été calculé après dénombrement sur boîte des bactéries ayant adhéré après 3h30 sur plaque polystyrène recouverte de mucine (A) et après 48h sur plaque polystyrène (B).

### Détermination du motif de liaison à l'ADN du RR0219.

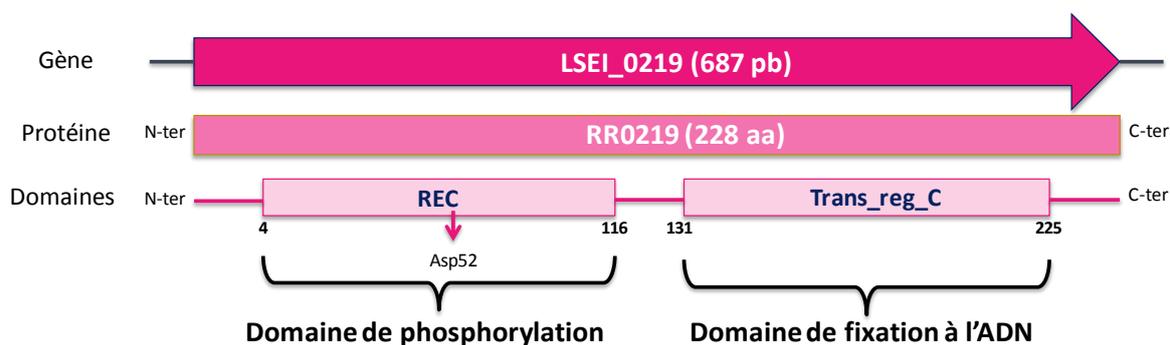
La protéine RR0219 purifiée a été utilisée pour réaliser des EMSAs. La capacité de phosphorylation avec de l'acétyl-phosphate a été vérifiée après chaque purification par migration sur gel BN-PAGE comme montré dans la Figure 35. Plusieurs tampons de fixation, de migration, et plusieurs types de gel ont été testés sans pour autant obtenir un retard avec la plus grande sonde (F-189R254 = Prom0219). L'hypothèse émise, à ce moment là, était que l'étiquette His<sub>6</sub> en C-terminal de la protéine pouvait empêcher la fixation du régulateur sur la sonde par un encombrement stérique car le domaine effecteur de la protéine se situe en C-terminal (Figure 36). Un nouveau clonage a alors été réalisé dans le pET28a<sup>+</sup> afin de produire une protéine avec une étiquette His<sub>6</sub> en N-terminal de la protéine qui permet ensuite le clivage de l'étiquette His<sub>6</sub> grâce au site reconnu par la thrombine présent dans le pET28a<sup>+</sup> et codé dans la protéine. Pour cela, le pTZ0219 (pTZ19R avec clonage du gène LSEI\_0219 au niveau du site de restriction SmaI) a été construit et utilisé pour construire le pET0219N. La phosphorylation de la protéine ainsi produite et clivée nommée RR0219N a aussi été vérifiée

et des EMSAs réalisés. Cependant, aucun retard n'a pu être observé malgré les différentes conditions testées. La protéine a peut-être besoin d'un cofacteur (protéine ou ion) pour se fixer à l'ADN ou ne se replie pas dans une bonne conformation lors de l'expression hétérologue et ne peut se fixer. Des EMSAs ont alors été réalisés avec les extraits protéiques cytoplasmiques de la souche parentale et du mutant Δ0219 (voir article précédent). Puis, pour déterminer le motif de fixation, la taille de la sonde a été diminuée et des EMSAs ont été réalisés avec l'extrait protéique cytoplasmique de *L. casei* ATCC 334 (Figure 37). Deux sondes de taille réduite, correspondant à la première partie (F-189R-1) et à la deuxième partie (F1R254) de la sonde Prom 0219 (F-189R254) ont été réalisées (voir Figure 32 du matériels et méthodes). Un retard a été obtenu avec les deux nouvelles sondes. Plusieurs motifs prédictifs ont été identifiés, ils sont présentés dans la Figure 38.

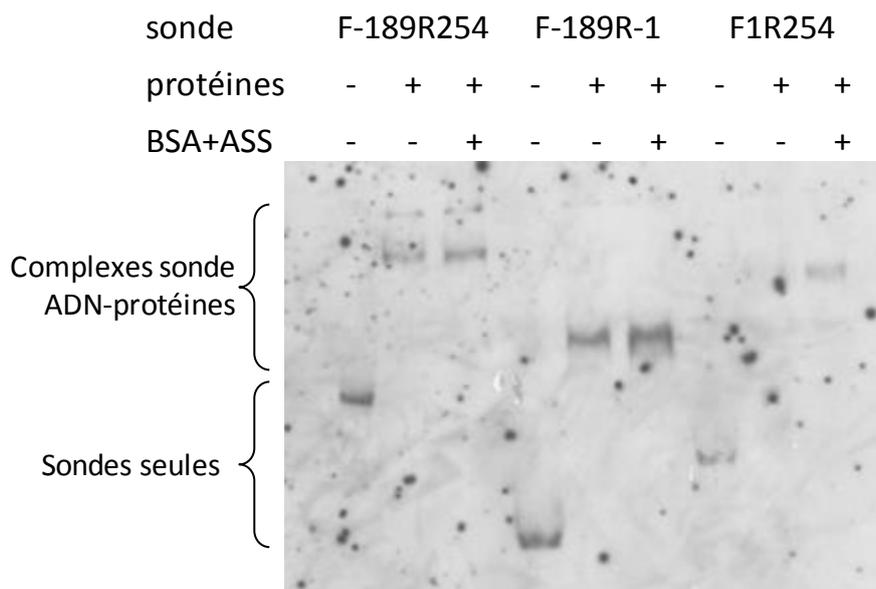


**Figure 35 : Phosphorylation du RR0219 en présence d'acétyl-phosphate.**

Les produits de phosphorylation ont été analysés par BN-PAGE 12,5% après différents temps d'incubation en présence d'acétyl-phosphate (60mM). 5 µg de RR0219 a été utilisé pour chaque réaction.



**Figure 36 : Représentation schématique du gène LSEI\_0219, de la protéine RR0219 et des différents domaines la constituant.**



**Figure 37 : EMSA de l'extrait protéique cytoplasmique de *L. casei* ATCC 334 avec trois sondes ADN représentant la région intergénique des gènes LSEI\_0218 et LSEI\_0219.**

Le retard a été réalisé avec 20 µg d'extrait protéique cytoplasmique en absence ou en présence de 2,5 µg.ml<sup>-1</sup> de BSA et ASS. BSA, sérum albumine bovine ; ASS, ADN de sperme de saumon.

```

5' GCTAAGGCTTTTGTCAATATGGTTATTGAGTCGTGGCGGGCTGATGATGCTGGTAAGACAACCTGTAGCTG
  ACGAACACCTCGTAACCAATGGGGATGTGTGGCAAGGCGTAAGGTGAGCCATTGGGAGGATTGCCTTAATA
  GCCAGTAAGTACTTCCGCTTGAGCGGGCAATGTTAATGCCAGCGATGTGCATGGTTGAGGGTAAACTTGCT
  GTGACCATTTTTGTCAAATCTTTCAATTAGCCGCCCTTCTTCGCGATTAAATCAACACTTTTGTACCTATT
  CTAGCCCTGTTTTCTTAGGCTTTTCGTAAAGAATTCCCCGGTCAACTTTGGTACAATGGACTTTGATGTAA
  CAGACCCTACAGTTCTTCTCATAAACGGAGCCAATACAATGAAAATTTTAATTGTTGACGACGATAAAGAA
  ATCGTCGAATTACTAAGT 3'
  
```

**Figure 38 : Représentation des motifs putatifs de liaison à l'ADN du RR0219.**

La séquence représentée correspond à la sonde Prom0219, la séquence en noire représente la séquence de la sonde F-189R-1 et la séquence en bleu représente la séquence de la sonde F1R254. Les nucléotides surlignés en gris correspondent aux codons initiateurs de traduction des ARNm des gènes LSEI\_0218 et LSEI\_0219. Les flèches de même couleur représentent les différents domaines de fixation possible. Les nucléotides soulignés correspondent soit à ceux qui se répètent soit à ceux qui sont inversés répétés.

### Induction du système à deux composants RR0219-HK0220

Pour déterminer, les conditions d'induction du système à deux composants, le taux de transcrits du gène LSEI\_0219 a été évalué chez la souche parentale en phase exponentielle de croissance dans différentes conditions en comparaison au taux de transcrits en condition témoin incubée à 37°C dans du MRS ou du tampon phosphate. Les ARNs de la souche parentale ont été extraits après 1 h d'incubation dans les différents milieux. Les différentes conditions testées sont les suivantes :

- NaCl : MRS + NaCl 1M
- Pénicilline : MRS + Pénicilline 0,1 µg.ml<sup>-1</sup>
- Vancomycine : MRS + Vancomycine 0.5 g.l<sup>-1</sup>
- Bile : MRS + Bile 3 g.l<sup>-1</sup>

- EtOH : MRS + Ethanol 12% (v/v)
- HCl : MRS acidifié avec de l'HCl à pH 3.0
- Ac lactique: MRS acidifié avec de l'acide lactique à pH 3.0
- Sans Glc : MRS sans glucose
- PPB 4°C : tampon phosphate maintenu au bain marie à 4°C
- PPB 42°C : tampon phosphate maintenu au bain marie à 42°C
- PPB 100X : tampon phosphate dans lequel les bactéries sont concentrées 100 fois par rapport au témoin
- PPB SDS : tampon phosphate + 0,05% SDS (m/v).

Le taux de transcrits relatif du gène LSEI\_0219 de *L. casei* dans les différentes conditions est présenté dans la Figure 39. Au vu des écart-types obtenus entre les différentes répétitions biologiques, le temps de contact d'une heure est probablement trop long et devrait être réduit afin de diminuer les variations d'expression obtenues. Néanmoins, les résultats obtenus indiquent une augmentation d'expression en présence de NaCl (stress osmotique), de bile et d'éthanol, en milieu acide (HCl et acide lactique) et lorsque la température est à 42°C. Ces conditions peuvent toutes induire une perturbation au niveau de la membrane cellulaire. Cette hypothèse a alors été vérifiée par mesure de la fluidité membranaire après ajout de bile, d'éthanol et augmentation de la température (Figure 40). Une modification de la fluidité membranaire mesurée par anisotropie de fluorescence de la sonde DPH (voir partie Matériels et Méthodes) a été observée dans les trois conditions testées. Alors que la bile et l'éthanol semblent fluidifier la membrane de manière transitoire, une température de 55°C induit une fluidification suivie d'une rigidification de la membrane.

L'induction du système à deux composants serait donc une modification de la fluidité membranaire, ce résultat coïncide avec la structure du senseur qui d'après l'analyse *in silico* posséderait un domaine senseur de 25 résidus d'acides aminés et serait capable de détecter des signaux liés au stress de l'enveloppe cellulaire (Mascher 2006; Mascher *et al.* 2006). Cependant le contexte génomique ne permet pas de classer l'HK0220 ni dans la catégorie des HKs BceS-like qui sont fonctionnellement et génétiquement liées à des ABC transporters ni dans la catégorie des HKs LiaS-like qui font parties d'un système à trois composants, le troisième composant étant une protéine membranaire agissant comme régulateur du système à deux composants en empêchant le transfert du phosphate de l'HK vers son RR.

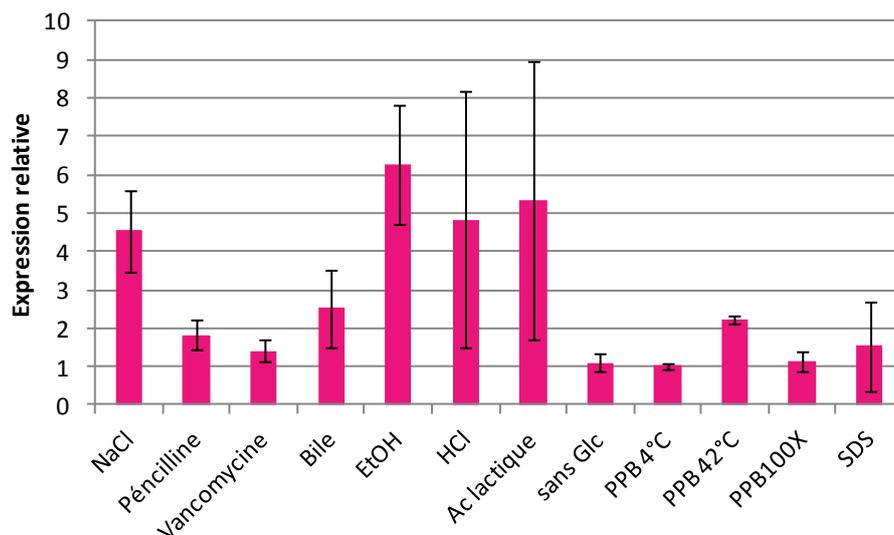


Figure 39 : Taux de transcrits du gène LSEI\_0219 obtenu après 1h dans les différentes conditions testées.

Les barres représentent l'écart-type entre les trois répétitions biologiques indépendantes.

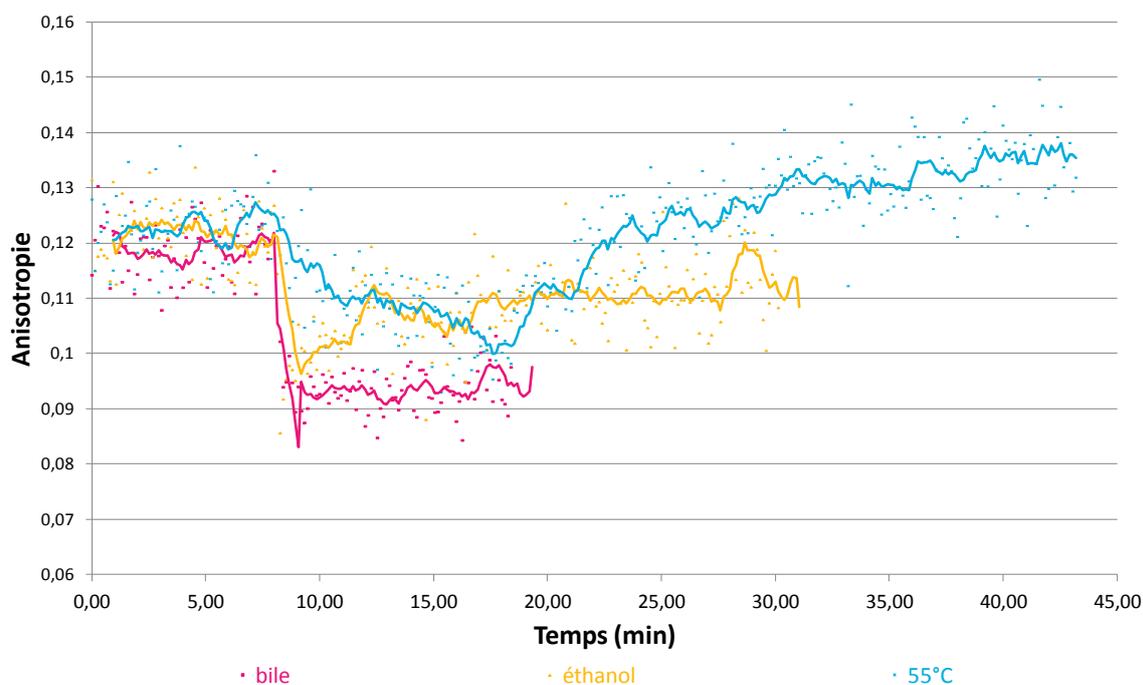


Figure 40 : Fluidité membranaire de *L. casei* en différentes conditions par mesure d'anisotropie de fluorescence.

Les cellules marquées au DPH (2µM) ont été mises en contact, à partir du temps 8 min, avec 0,003% de bile ou 12% d'éthanol ou soumis à une augmentation de la température à 55°C.

#### Expériences complémentaires à réaliser pour finaliser l'article 4

La compétition *ex vivo* en lumen de lapin pour les trois mutants par rapport au sauvage doit être testée de la même façon que la bile, l'acide et le lysozyme.

L'extraction des protéines de surface des mutants et de la souche parentale doit être réalisée en phase exponentielle de croissance afin de déterminer si la présentation des deux antigènes

de surface est aussi affectée à ce stade de croissance et de voir si les résultats de transcription à ce stade de croissance sont corrélés, donc si la modulation de la quantité des Ag de surface provient d'une régulation transcriptionnelle.

L'ajout de senseur pour réaliser les EMSAs à partir de la protéine RR0219 purifiée peut être testé. Le senseur peut peut-être aider à la fixation ou au bon repliement du régulateur (Bisicchia *et al.* 2007).

# **Conclusion Générale**

## La mutagenèse par transposition

Différentes approches globales existent pour étudier les gènes bactériens impliqués dans la réponse à des stress ou l'adaptation à des milieux ou écosystèmes divers. Par exemple les études post-génomique : études transcriptomiques et protéomiques. L'étude transcriptomique peut être réalisée par la méthode des puces à ADN qui permettent d'analyser le niveau de transcrit correspondant aux ADN cibles ordonnés sur la puce (généralement l'ensemble des gènes du génome) dans une condition donnée. L'étude protéomique permet quant à elle d'identifier et de quantifier les variations du taux d'expression de protéines dans différentes conditions. La limite de ces types d'approche provient du fait que le criblage des gènes est réalisé comparativement à une condition témoin (généralement une culture *in vitro*).

La stratégie d'étude que nous avons utilisée pour la recherche globale des gènes impliqués dans l'établissement de *L. casei* dans l'intestin est la mutagenèse aléatoire associée à une stratégie de criblage approprié. La méthode la plus appropriée pour la stratégie d'étude et la plus informative était la mutagenèse par transposition. Cette méthode permettait de caractériser les mutants avec un phénotype d'intérêt directement après criblage et d'éviter de réaliser une mutagenèse ciblée sur les gènes d'intérêt par la suite. La méthode de mutagenèse par transposition utilisée sur *L. casei* a la particularité d'utiliser deux vecteurs, le premier codant pour une transposase et le second possédant les cibles d'insertion de transposase et correspondant ainsi au transposon (Article 1). Cette méthode de mutagenèse a été mise au point et utilisée pour pallier le faible taux de transformation de la souche et peut être utilisé pour d'autres bactéries lactiques. L'obtention de la banque de mutant constituait une étape clé pour la réalisation de cette thèse.

## L'identification des sites d'insertion du transposon des mutants

Pour rendre la banque de mutants doublement informative (identification des mutants avec phénotype mais aussi des mutants sans phénotype) et permettre de pré-sélectionner des mutants avant le criblage, le séquençage de la région d'insertion du transposon a été réalisé. Ceci a été rendu possible grâce à l'optimisation (i) de l'extraction d'ADN en plaque 96 puits pour une bactérie difficile à lyser comme *L. casei* et (ii) du protocole de séquençage par la technique de Sanger (Article 2). Ainsi, 87% des mutants ont pu être identifiés. La banque ainsi annotée peut être utilisée pour différents criblages fonctionnels et constitue un outil d'intérêt pour l'équipe.

## Le criblage fonctionnel des mutants quant à leur capacité à s'établir dans l'intestin

La Signature Tagged Mutagenesis (STM) a été développée à partir de la mutagenèse par transposition afin de réaliser le criblage en ensemble de mutants et par conséquent de limiter le nombre d'animaux utilisés pour l'étude *in vivo*. Grâce au séquençage de l'insertion du transposon des mutants de la banque, les mutants ayant un gène interrompu ont pu être sélectionnés et réorganisés pour créer une banque de mutants ayant le transposon d'intégré dans la séquence d'un gène. Le criblage de cette banque annotée *in vivo* permettait d'identifier sans *a priori* les gènes impliqués dans l'établissement de *L. casei* dans l'intestin mais aussi ceux qui ne le sont pas. Ainsi 47 gènes ont été identifiés, ils encodent pour des protéines de fonctions variées : les processus cellulaires basiques, le métabolisme, la biosynthèse de la paroi cellulaire et l'adaptation au milieu environnant. Les protéines identifiées ont une portée plus générale que la simple réponse au stress classique telle que le repliement des protéines (augmentation de la quantité de heat ou cold shock protein...) ou la réparation de l'ADN et sont davantage impliquées dans la physiologie, l'adaptation de la bactérie. Le milieu intestinal n'est pas perçu comme un stress fort pour cette bactérie. La limite du criblage réalisé en ensemble de mutants est que les mutants (pour des protéines de surface par exemple) complémentés par d'autres mutants ne pouvaient être identifiés. Pour ces mutants là, il faudrait faire un autre criblage où les mutants seraient cette fois testés individuellement sur culture cellulaire par exemple.

## Caractérisation fonctionnelles de protéines impliquées dans l'établissement dans l'intestin

### *Choix des gènes à caractériser*

A la suite du criblage *in vivo*, plusieurs gènes pouvaient être caractérisés, le choix s'est porté sur l'opéron LSEI\_0219-0221 pour trois raisons principales. Premièrement, les trois mutants ( $\Delta 0219$ ,  $\Delta 0220$  et  $\Delta 0221$ ) de chaque gène avaient des difficultés à s'établir *in vivo*. Deuxièmement, le gène LSEI\_0221 code pour une enzyme intervenant dans la biosynthèse du peptidoglycane (PG), le PG étant le constituant majeur de la paroi cellulaire et cette paroi étant le premier site de reconnaissance/contact entre l'hôte ou le milieu et la bactérie. Et troisièmement, les deux autres gènes de l'opéron codent pour un système à deux composants (TCS). Ce TCS était le seul des 16 TCS complets de *L. casei* (voir synthèse bibliographique) à afficher un phénotype *in vivo* d'après notre criblage (au moins un mutant pour le gène codant soit pour le RR soit pour l'HK pour 10 des 16 TCS).

De plus, l'enzyme codée par LSEI\_0221, intervenant dans la biosynthèse du PG est une des 13 peptidoglycan hydrolases (PGH) (LSEI\_0020, 0221, 0236, 0281, 0393, 0539, 0599, 1072, 1116, 1536, 1909, 2029 et 2226) qui existent d'après l'annotation du génome. Sur ces 13 gènes, 9 mutants ont été obtenus (gènes soulignés ci-dessus). Le mutant pour le gène LSEI\_0281 avait un phénotype *in vivo* différent mais il a été retiré car la stratégie de criblage ne permettait pas de le discriminer. En effet, la morphologie de ce mutant était très différente de la souche parentale (*L. casei* ATCC 334). Cette morphologie en filament très long engendrait un défaut de corrélation DO/UFC empêchant ainsi de savoir si le phénotype *in vivo* était dû à l'incapacité de s'établir *in vivo* ou seulement la conséquence de la morphologie du mutant. Le mutant pour le gène LSEI\_0221 a donc été considéré comme le seul mutant pour une PGH à avoir un phénotype différent *in vivo*. La position des gènes codant pour le TCS (en amont du gène LSEI\_0221) laissait supposer une implication de ce TCS dans la biosynthèse du PG. Les protéines codées par les gènes de l'opéron ont été caractérisées fonctionnellement puis l'implication de ces protéines dans l'établissement dans l'intestin a été recherchée.

#### Caractérisation du système à deux composants

La fonctionnalité de transfert de phosphate de la HK0220 (histidine kinase codée par le gène LSEI\_0220) vers le RR0219 (response regulator codé par le gène LSEI\_0219) du TCS a été vérifiée (voir synthèse bibliographique pour la définition des fonctions protéiques citées et article 4 pour la vérification de la fonction). Le RR0219 annoté comme appartenant à la famille d'OmpR possède un domaine effecteur de liaison à l'ADN (voir synthèse bibliographique). Malgré diverses tentatives infructueuses de retard sur gel avec une sonde couvrant la région intergénique entre les gènes LSEI\_0218 et LSEI\_0219, aucun retard avec le RR0219 purifié phosphorylé ou non n'a pu être observé. Le mauvais repliement de la protéine chez *E. coli* ou la nécessité d'un élément cellulaire, et/ou d'un cofacteur peut être une hypothèse expliquant l'absence de fixation de la protéine sur la sonde. Les retards sur gel ont alors été réalisés avec l'extrait protéique cellulaire de la souche parentale et celui du mutant  $\Delta$ 0219 ( $\Delta$ LSEI\_0219). L'extrait protéique cellulaire de la souche parentale a permis la formation d'un complexe avec la sonde alors que celui du mutant ne le permet pas. Ce résultat laisse supposer l'implication du RR0219 dans la formation du complexe. D'autres retards sur gel ont été réalisés avec deux sondes correspondant à la moitié de la plus grande sonde, un retard avec l'extrait protéique cellulaire de la souche parentale a également été obtenu avec ces deux plus petites sondes. L'analyse *in silico* des sondes a montré différents sites potentiels de fixation qui demandent à présent à être confirmés. Le signal perçu par le domaine senseur (voir synthèse bibliographique) de l'HK0220 reste quant à lui encore inconnu, cependant, d'après les résultats de transcription et d'anisotropie de fluorescence, une perturbation membranaire semble induire le système. L'identification du motif reconnu par le RR0219

permettra de déterminer quels gènes sont potentiellement régulés et la prédiction des fonctions des protéines codées permettra de réduire le champ des possibilités de signaux perçus.

Aucun mutant n'a été obtenu pour les gènes LSEI\_2808 et LSEI\_2807 codant pour le TCS RR2808-HK2807 qui semblerait lui aussi d'après l'annotation impliqué dans la biosynthèse de la paroi. Il serait intéressant d'obtenir un mutant pour ce TCS et de le tester quant à sa capacité à s'établir dans l'intestin et de caractériser sa fonction.

#### *Caractérisation fonctionnelle de la PBP0221 (Figure 41)*

La caractérisation fonctionnelle de l'enzyme impliquée dans la biosynthèse du PG, la PBP0221 (penicillin-binding protein codée par le gène LSEI\_0221) a montré une fonction différente de celle attendue d'après l'annotation du génome. La PBP0221 aurait plutôt une fonction de L,D-carboxypeptidase que de D-Alanyl-D-Alanine carboxypeptidase (voir article 4). La fonction de D-Alanyl-D-Alanine carboxypeptidase est nécessaire pour permettre l'assemblage du PG puisqu'elle enlève le dernier résidu du pentapeptide des brins de PG permettant la transpeptidation (voir Fig. 23 de la partie Synthèse Bibliographique). Aucun autre gène n'est annoté de cette façon dans le génome de *L. casei* (Cai *et al.* 2009) cependant la recherche *in silico* de gènes potentiellement impliqués dans cette fonction permet de proposer plusieurs candidats tels que les gènes annotés comme étant des « Beta-lactamase class C and other penicillin-binding proteins » (LSEI\_1152, 0435, 1802, 1805, 2137, et 2773) (Figure 41). Le gène LSEI\_2553 annoté comme hypothetic protein (impliqué aussi dans l'établissement *in vivo*) pourrait lui aussi avoir cette fonction, en effet, la séquence protéique codée par ce gène montre 63 % d'identité avec une protéine annotée comme D-alanyl-D-alanine carboxypeptidase chez *L. rhamnosus* K32 (Figure 41).

#### *Rôle de l'opéron dans la surface cellulaire (Figure 41)*

L'absence fonctionnelle de la PBP0221 induit une modification du PG, cette modification est viable pour la bactérie en condition *in vitro* mais pas en condition *in vivo* dans l'intestin.

L'analyse transcriptionnelle des gènes impliqués dans la biosynthèse du PG a montré une diminution de la transcription en phase stationnaire plus importante pour trois gènes (LSEI\_0020, 2029 et 2226) codant tous les trois pour des PGHs (voir article 4). Alors que la transcription de ces trois gènes est diminuée, la transcription du gène LSEI\_0539, codant aussi pour une PGH, est quant à elle augmentée chez le mutant  $\Delta 0221$  sans pour autant modifier la structure du PG du mutant  $\Delta 0221$  par rapport aux autres mutants ( $\Delta 0219$  et

Δ0220). L'augmentation de transcription du gène LSEI\_0539 semble être régulée par le TCS (pas d'augmentation dans les mutants du TCS). Cependant l'analyse des protéines de surface a montré une diminution de la présentation uniquement des deux antigènes de surface (AgS) codés par les gènes LSEI\_0020 et LSEI\_2029. L'augmentation de transcription du gène LSEI\_0539 pour le mutant Δ0221 n'engendre pas un profil de protéines de surface différent, la protéine est peut être sécrétée ou bien ancrée dans la membrane ou encore liée de manière covalente au peptidoglycane.

La modification de présentation des AgS peut être due à la composition du PG différente de la souche parentale (problème d'ancrage) et/ou seulement à la régulation transcriptionnelle en réponse à l'absence de la PBP0221. L'absence de l'AgS codé par le gène LSEI\_2029 ne peut expliquer à elle seule l'incapacité des mutants à s'établir dans l'intestin car le mutant pour ce gène n'a pas de phénotype *in vivo*. Il serait intéressant de déterminer si, chez ce mutant, la présentation de l'AgS codé par le gène LSEI\_0020 est augmentée pour compenser l'absence. Le mutant pour le gène LSEI\_0020 n'a pas été obtenu dans la banque de mutant, il est donc difficile de statuer quant à la nécessité de la protéine dans l'établissement dans l'intestin. Cependant le gène LSEI\_0020 n'est pas indispensable à cette espèce car un mutant pour ce gène a été obtenu chez *L. casei* BL23 (Bauerl *et al.* 2010) mais est indispensable à *L. rhamnosus* GG car aucun mutant n'a pu être obtenu (Claes *et al.* 2012). La surexpression d'un ou des deux AgS de *L. casei* pourrait être réalisée chez la souche parentale et la capacité à s'établir dans l'intestin pourrait être testée.

Les protéines permettant d'hydrolyser les liens peptidiques du PG (endopeptidases et *N*-acetylmuramoyl-L-alanine amidases) d'après les domaines les constituants (Layec *et al.* 2008) sont codées, chez *L. casei*, par les gènes LSEI\_0020, 0281, 0539, 2029 et 1536 (Figure 41). Chez *L. rhamnosus* GG toutes ces protéines sont présentes à l'exception de celle codée par le gène LSEI\_0539 ; l'absence de cette protéine peut peut-être expliquer pourquoi aucun mutant pour le gène correspondant au gène LSEI\_0020 n'a pu être obtenu chez *L. rhamnosus* GG. Chez le mutant Δ0221, l'augmentation d'expression du gène LSEI\_0539, potentiellement régulé par le TCS, a peut-être finalement lieu pour compenser la diminution d'expression du gène LSEI\_0020 et pas l'absence fonctionnelle de la PBP0221.

La diminution de transcription de nombreuses enzymes impliquées dans la synthèse du PG et la sensibilité augmentée des mutants à la bacitracine laisse supposer un renouvellement moins important du PG. Le suivi de l'incorporation de [<sup>3</sup>H]GlcNAc dans le PG des mutants pourrait être testé en comparaison à celle de la souche parentale et permettrait de déterminer si l'absence de la PBP0221 engendre un turnover du PG diminué.

La composition différente du PG des mutants peut aussi moduler l'ancrage des protéines liées de manières covalentes ou non au PG (voir synthèse bibliographique pour les types de protéines ancrées). Une analyse transcriptionnelle pourrait être réalisée pour étayer cette hypothèse.

*Rôle de l'opéron dans l'interaction avec l'hôte*

Les mutants  $\Delta 0219$ ,  $\Delta 0220$  et  $\Delta 0221$  ont montré une sensibilité aux peptides antimicrobiens cationiques hBD1, hBD2 et CCL20 alors que la souche parentale n'est pas sensible. Cette différence peut expliquer à elle seule la diminution de quantité des mutants lors du séjour dans l'intestin.

La modification du peptidoglycane et la diminution de présentation d'AgS peuvent engendrer (i) au niveau de la cellule, une faiblesse induisant ainsi un manque de résistance de la bactérie à des conditions physicochimiques et (ii) au niveau de l'interaction hôte-bactérie, une reconnaissance différente de l'hôte induisant alors une réponse différente (peptides antimicrobiens cationiques). La composition de la surface cellulaire induit-elle une différence de réponse de l'hôte (hypothèse n°1) ou est-ce la bactérie qui ne peut pas faire face, de par sa composition de la surface cellulaire différente, à la réponse basale de l'hôte (hypothèse n°2)? Le criblage qui a été conduit a permis de valider l'hypothèse n°2 alors qu'il n'a pas permis de statuer quant à l'hypothèse n°1. En effet, si une protéine de surface (ou sécrétée), nécessaire pour induire ou réprimer une réponse de l'hôte, est absente chez un mutant, la modulation de la réponse de l'hôte est néanmoins réalisée par cette protéine de surface présente (ou sécrétée) chez les autres mutants. L'hypothèse n°1, qui n'a pas été expérimentalement testée, n'est cependant pas exclue puisque les deux hypothèses peuvent coexister.

Une sensibilité plus importante du mutant  $\Delta 0221$  est observée pour les défensines hBD1 et hBD2 par rapport aux autres mutants, cette augmentation de sensibilité peut être due à l'augmentation de la présence de la PGH0539 (protéine codée par le gène LSEI\_0539 surexprimé chez ce mutant) ou à la diminution de présentation des 16 autres protéines impliquées dans la biosynthèse du PG ayant potentiellement une régulation transcriptionnelle contrôlée par le TCS.

Les PGHs peuvent induire à elles seules une réponse cellulaire des cellules épithéliales, c'est le cas, par exemple, des protéines p40 et p75 chez *L. rhamnosus* GG (Seth *et al.* 2008; Yan *et al.* 2007) et des orthologues de *L. casei* ATCC 334 codés respectivement par les gènes LSEI\_0020 et LSEI\_0281. De la même façon, la réponse cellulaire des cellules épithéliales pourrait être déterminée en présence des protéines PGH0539, AgS2029 et PGH2226 codées respectivement par les gènes LSEI\_0539, LSEI\_2029 et LSEI\_2226. L'activité catalytique des protéines AgS0020, PGH0539, AgS2029 et PGH2226 codées respectivement par les gènes LSEI\_0020, 0539, 2029 et 2226 pourrait être déterminée afin de savoir quels types de molécules constituant le PG peuvent se retrouver dans le milieu extracellulaire de la bactérie et seraient susceptibles d'être reconnus par les cellules constituant l'épithélium intestinal comme c'est le cas par exemple du muramyl-dipeptide (NAG-NAM-dipeptide) (Girardin *et al.* 2003; Nigro *et al.* 2014).

Ce travail de thèse a permis d'obtenir une banque de mutants de *L. casei* annotée. L'annotation d'une banque est un réel atout car elle permet une approche nouvelle du criblage fonctionnel ; en effet, l'annotation permet de prendre en compte l'intégralité des gènes qui sont testés en opposant les gènes qui sont impliqués dans un phénotype particulier et ceux qui ne le sont pas. L'annotation donne également la possibilité d'éliminer les mutants redondants et de s'affranchir de la nécessité de la mutagenèse ciblée (par recombinaison homologue) après un criblage pour caractériser les mutants d'intérêt. La banque de mutant annotée peut à présent être utilisée pour de nombreux criblages avec diverses retombées à la fois fondamentales et appliquées. A la suite du criblage dans l'anse iléale ligaturée de lapin, nous avons pu démontrer qu'une architecture convenable de la surface cellulaire bactérienne est indispensable à la survie dans l'intestin et plus particulièrement que les hydrolases du peptidoglycane (PGH) sont nécessaires à la protection de la bactérie dans l'environnement intestinal. Parmi les 44 autres protéines identifiées d'après le criblage, de nombreuses méritent d'être caractérisées afin de trouver d'autres voies d'implication originales dans l'établissement de *L. casei* dans l'intestin.

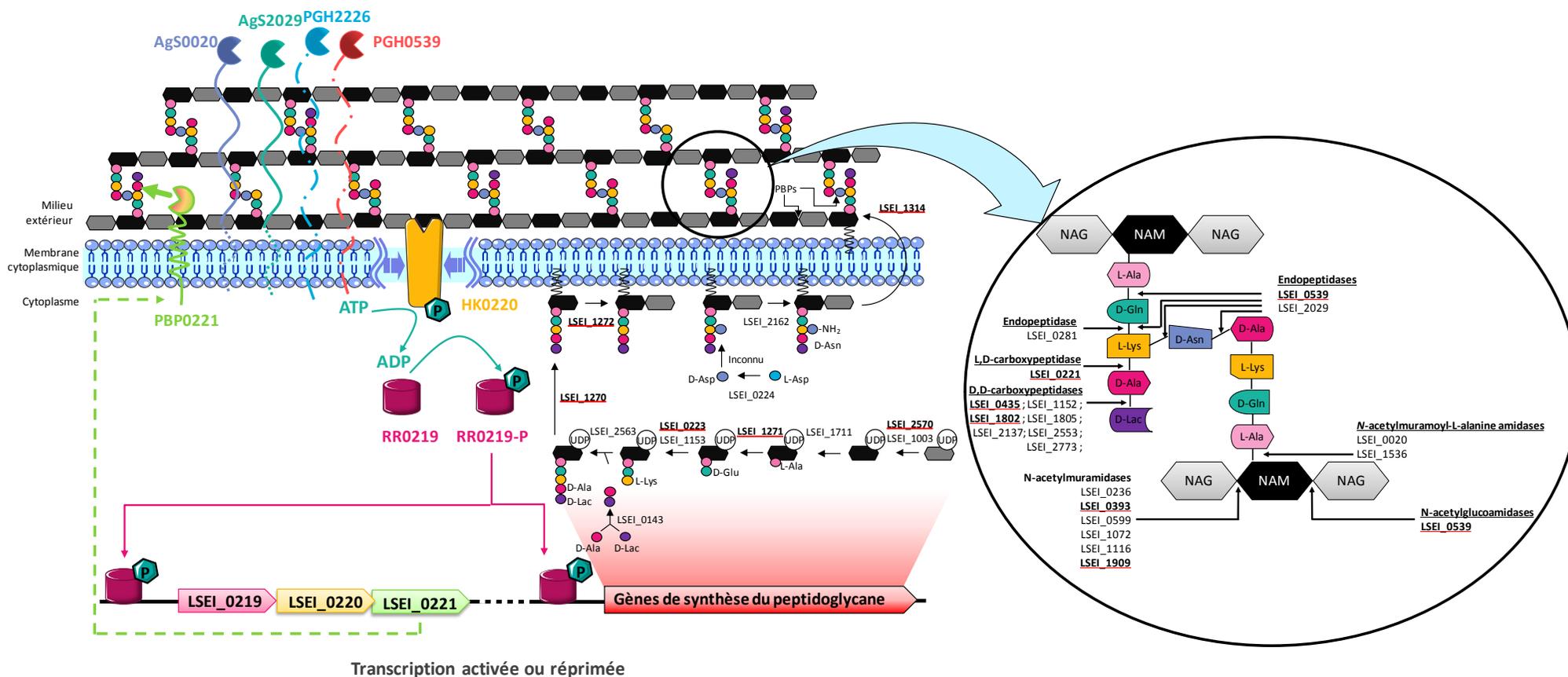


Figure 41 : Rôles démontrés et probables de l'opéron LSEI\_0219-0221 dans la surface cellulaire de *L. casei*.

L'HK0220 serait capable de détecter un changement membranaire. Elle est capable de s'autophosphoryler et de transférer son phosphate au RR0219. Le RR0219 permettrait d'activer ou de réprimer la transcription de son propre opéron et de gènes intervenant dans la synthèse du peptidoglycane. Le nom des gènes correspond aux protéines impliquées dans la réaction indiquée par les flèches noires. Les gènes en gras et soulignés en rouge correspondent à ceux qui sont potentiellement régulés par le TCS RR0219-HK0220. Les protéines de surface AgS0020, AgS2029, PGH2226 et PGH0539 sont celles dont la transcription varie le plus chez les mutants de l'opéron. Les protéines codées par les gènes LSEI\_0281 et LSEI\_0221 ont des fonctions expérimentalement démontrées. La fonction des autres protéines a été attribuée selon l'annotation du génome excepté pour la catégorie des DD carboxy dont la fonction a été déterminée par analyse in silico (voir texte). Inconnu indique qu'aucun gène annoté pour une protéine ayant cette fonction n'a été trouvé dans le génome. RR0219 : Response Regulator codé par le gène LSEI\_0219 ; RR0219-P : RR0219 phosphorylé ; HK0220 : Histidine Kinase codée par le gène LSEI\_0220 ; PBP0221 : Penicillin-Binding Protein ou L,D-carboxypeptidase codée par le gène LSEI\_0221 ; AgS0020 : Surface Antigen codé par le gène LSEI\_0020 ; AgS2029 : Surface Antigen codé par le gène LSEI\_2029 ; PGH2226 : Peptidoglycan hydrolase codée par le gène LSEI\_2226 ; PGH0539 : Peptidoglycan hydrolase codée par le gène LSEI\_0539.

**Références**  
**Bibliographiques**

1. **Abachin E, Poyart C, Pellegrini E, Milohanic E, Fiedler F, Berche P, Trieu-Cuot P.** 2002. Formation of D-alanyl-lipoteichoic acid is required for adhesion and virulence of *Listeria monocytogenes*. *Mol Microbiol* **43**(1):1-14
2. **Aguilar PS, Hernandez-Arriaga AM, Cybulski LE, Erazo AC, de Mendoza D.** 2001. Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilis*. *EMBO J* **20**(7):1681-1691
3. **Alcantara C, Revilla-Guarinos A, Zuniga M.** 2011. Influence of two-component signal transduction systems of *Lactobacillus casei* BL23 on tolerance to stress conditions. *Appl Environ Microbiol* **77**(4):1516-1519
4. **Alexander RP, Lowenthal AC, Harshey RM, Ottemann KM.** 2010. CheV: CheW-like coupling proteins at the core of the chemotaxis signaling network. *Trends Microbiol* **18**(11):494-503
5. **Araki Y, Oba S, Araki S, Ito E.** 1980. Enzymatic deacetylation of *N*-acetylglucosamine residues in cell wall peptidoglycan. *J Biochem* **88**(2):469-479
6. **Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM.** 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* **29**(2):231-262
7. **Archibald AR.** 1985. Structure and assembly of the cell wall in *Bacillus subtilis*. *Biochem Soc Trans* **13**(6):990-992
8. **Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Dore J, Antolin M, Artiguenave F, Blottiere HM, Almeida M, Brechot C, Cara C, Chervaux C, Cultrone A, Delorme C, Denariáz G, Dervyn R, Foerstner KU, Friss C, van de Guchte M, Guedon E, Haimet F, Huber W, van Hylckama-Vlieg J, Jamet A, Juste C, Kaci G, Knol J, Lakhdari O, Layec S, Le Roux K, Maguin E, Merieux A, Melo Minardi R, M'Rini C, Muller J, Oozeer R, Parkhill J, Renault P, Rescigno M, Sanchez N, Sunagawa S, Torrejon A, Turner K, Vandemeulebrouck G, Varela E, Winogradsky Y, Zeller G, Weissenbach J, Ehrlich SD, Bork P.** 2011. Enterotypes of the human gut microbiome. *Nature* **473**(7346):174-180
9. **Asong J, Wolfert MA, Maiti KK, Miller D, Boons GJ.** 2009. Binding and Cellular Activation Studies Reveal That Toll-like Receptor 2 Can Differentially Recognize Peptidoglycan from Gram-positive and Gram-negative Bacteria. *J Biol Chem* **284**(13):8643-8653
10. **Avall-Jaaskelainen S, Palva A.** 2005. *Lactobacillus* surface layers and their applications. *FEMS Microbiol Rev* **29**(3):511-529
11. **Baikalov I, Schroder I, Kaczor-Grzeskowiak M, Grzeskowiak K, Gunsalus RP, Dickerson RE.** 1996. Structure of the *Escherichia coli* response regulator NarL. *Biochemistry* **35**(34):11053-11061
12. **Bauerl C, Perez-Martinez G, Yan F, Polk DB, Monedero V.** 2010. Functional analysis of the p40 and p75 proteins from *Lactobacillus casei* BL23. *J Mol Microbiol Biotechnol* **19**(4):231-241
13. **Bellon-Fontaine M-N, Rault J, van Oss CJ.** 1996. Microbial adhesion to solvents: a novel method to determine the electron-donor/electron-acceptor or Lewis acid-base properties of microbial cells. *Colloids and Surf B: Biointerfaces* **7**:47-53
14. **Bernard E, Rolain T, Courtin P, Guillot A, Langella P, Hols P, Chapot-Chartier MP.** 2011. Characterization of *O*-acetylation of *N*-acetylglucosamine: a novel structural variation of bacterial peptidoglycan. *J Biol Chem* **286**(27):23950-23958
15. **Bertani G.** 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* **62**(3):293-300
16. **Bijlsma JJ, Burghout P, Kloosterman TG, Bootsma HJ, de Jong A, Hermans PW, Kuipers OP.** 2007. Development of genomic array footprinting for identification of

- conditionally essential genes in *Streptococcus pneumoniae*. Appl Environ Microbiol **73**(5):1514-1524
17. **Bisicchia P, Noone D, Lioliou E, Howell A, Quigley S, Jensen T, Jarmer H, Devine KM.** 2007. The essential YycFG two-component system controls cell wall metabolism in *Bacillus subtilis*. Mol Microbiol **65**(1):180-200
  18. **Boekhorst J, de Been MW, Kleerebezem M, Siezen RJ.** 2005. Genome-wide detection and analysis of cell wall-bound proteins with LPxTG-like sorting motifs. J Bacteriol **187**(14):4928-4934
  19. **Borrue N, Carol M, Casellas F, Antolin M, de Lara F, Espin E, Naval J, Guarner F, Malagelada JR.** 2002. Increased mucosal tumour necrosis factor alpha production in Crohn's disease can be downregulated ex vivo by probiotic bacteria. Gut **51**(5):659-664
  20. **Borrue N, Casellas F, Antolin M, Llopis M, Carol M, Espin E, Naval J, Guarner F, Malagelada JR.** 2003. Effects of nonpathogenic bacteria on cytokine secretion by human intestinal mucosa. Am J Gastroenterol **98**(4):865-870
  21. **Bourret RB.** 2010. Receiver domain structure and function in response regulator proteins. Curr Opin Microbiol **13**(2):142-149
  22. **Brinster S, Furlan S, Serror P.** 2007. C-terminal WxL domain mediates cell wall binding in *Enterococcus faecalis* and other gram-positive bacteria. J Bacteriol **189**(4):1244-1253
  23. **Broadbent JR, Larsen RL, Deibel V, Steele JL.** 2010. Physiological and transcriptional response of *Lactobacillus casei* ATCC 334 to acid stress. J Bacteriol **192**(9):2445-2458
  24. **Browning DF, Busby SJ.** 2004. The regulation of bacterial transcription initiation. Nat Rev Microbiol **2**(1):57-65
  25. **Buck BL, Altermann E, Svingerud T, Klaenhammer TR.** 2005. Functional analysis of putative adhesion factors in *Lactobacillus acidophilus* NCFM. Appl Environ Microbiol **71**(12):8344-8351
  26. **Buist G, Steen A, Kok J, Kuipers OP.** 2008. LysM, a widely distributed protein motif for binding to (peptido)glycans. Mol Microbiol **68**(4):838-847
  27. **Cai H, Thompson R, Budinich MF, Broadbent JR, Steele JL.** 2009. Genome sequence and comparative genome analysis of *Lactobacillus casei*: insights into their niche-associated evolution. Genome Biol Evol **1**:239-257
  28. **Call EK, Klaenhammer TR.** 2013. Relevance and application of sortase and sortase-dependent proteins in lactic acid bacteria. Front Microbiol **4**:73
  29. **Cao-Hoang L, Marechal PA, Le-Thanh M, Gervais P, Wache Y.** 2008. Fluorescent probes to evaluate the physiological state and activity of microbial biocatalysts: a guide for prokaryotic and eukaryotic investigation. Biotechnol J **3**(7):890-903
  30. **Cashin P, Goldsack L, Hall D, O'Toole R.** 2006. Contrasting signal transduction mechanisms in bacterial and eukaryotic gene transcription. FEMS Microbiol Lett **261**(2):155-164
  31. **Chaillou S, Champomier-Verges MC, Cornet M, Crutz-Le Coq AM, Dudez AM, Martin V, Beaufils S, Darbon-Rongere E, Bossy R, Loux V, Zagorec M.** 2005. The complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K. Nat Biotechnol **23**(12):1527-1533
  32. **Claes IJ, Schoofs G, Regulski K, Courtin P, Chapot-Chartier MP, Rolain T, Hols P, von Ossowski I, Reunanen J, de Vos WM, Palva A, Vanderleyden J, De Keersmaecker SC, Lebeer S.** 2012. Genetic and biochemical characterization of the cell wall hydrolase activity of the major secreted protein of *Lactobacillus rhamnosus* GG. PLoS One **7**(2):e31588
  33. **Collier-Hyams LS, Neish AS.** 2005. Innate immune relationship between commensal flora and the mammalian intestinal epithelium. Cell Mol Life Sci **62**(12):1339-1348
  34. **Cornett JB, Shockman GD.** 1978. Cellular lysis of *Streptococcus faecalis* induced with triton X-100. J Bacteriol **135**(1):153-160
  35. **Dagkessamanskaia A, Moscose M, Henard V, Guiral S, Overweg K, Reuter M, Martin B, Wells J, Claverys JP.** 2004. Interconnection of competence, stress and CiaR regulons in *Streptococcus pneumoniae*: competence triggers stationary phase autolysis of ciaR mutant cells. Mol Microbiol **51**(4):1071-1086

36. **Danne C, Dramsi S.** 2012. Pili of gram-positive bacteria: roles in host colonization. *Res Microbiol* **163**(9-10):645-658
37. **De Keersmaecker SC, Braeken K, Verhoeven TL, Perea Velez M, Lebeer S, Vanderleyden J, Hols P.** 2006. Flow cytometric testing of green fluorescent protein-tagged *Lactobacillus rhamnosus* GG for response to defensins. *Appl Environ Microbiol* **72**(7):4923-4930
38. **De Man PJ, Rogosa M, Sharpe. M.** 1960. A medium for the cultivation of *Lactobacilli*. *J Appl Bacteriol* **23**:130-135
39. **De Vuyst L, Degeest B.** 1999. Heteropolysaccharides from lactic acid bacteria. *FEMS Microbiol Rev* **23**(2):153-177
40. **Dego OK, Almeida RA, Oliver SP.** 2011. Presence of ISS1-like insertion sequence in wild type *Streptococcus uberis* strains isolated from cases of bovine mastitis. *Vet Microbiol* **151**(3-4):315-320
41. **Delcour J, Ferain T, Deghorain M, Palumbo E, Hols P.** 1999. The biosynthesis and functionality of the cell-wall of lactic acid bacteria. *Antonie Van Leeuwenhoek* **76**(1-4):159-184
42. **Dicks LM, Du Plessis EM, Dellaglio F, Lauer E.** 1996. Reclassification of *Lactobacillus casei* subsp. *casei* ATCC 393 and *Lactobacillus rhamnosus* ATCC 15820 as *Lactobacillus zaeae* nom. rev., designation of ATCC 334 as the neotype of *L. casei* subsp. *casei*, and rejection of the name *Lactobacillus paracasei*. *Int J Syst Bacteriol* **46**(1):337-340
43. **Dower WJ, Miller JF, Ragsdale CW.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* **16**(13):6127-6145
44. **Driessen AJ, Nouwen N.** 2008. Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem* **77**:643-667
45. **Dubrac S, Boneca IG, Poupel O, Msadek T.** 2007. New insights into the WalK/WalR (YycG/YycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in *Staphylococcus aureus*. *J Bacteriol* **189**(22):8257-8269
46. **Fabretti F, Theilacker C, Baldassarri L, Kaczynski Z, Kropec A, Holst O, Huebner J.** 2006. Alanine esters of enterococcal lipoteichoic acid play a role in biofilm formation and resistance to antimicrobial peptides. *Infect Immun* **74**(7):4164-4171
47. **Fisher SL, Jiang W, Wanner BL, Walsh CT.** 1995. Cross-talk between the histidine protein kinase VanS and the response regulator PhoB. Characterization and identification of a VanS domain that inhibits activation of PhoB. *J Biol Chem* **270**(39):23143-23149
48. **Francius G, Lebeer S, Alsteens D, Wildling L, Gruber HJ, Hols P, De Keersmaecker S, Vanderleyden J, Dufrene YF.** 2008. Detection, localization, and conformational analysis of single polysaccharide molecules on live bacteria. *ACS Nano* **2**(9):1921-1929
49. **Galperin MY.** 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J Bacteriol* **188**(12):4169-4182
50. **Galperin MY.** 2010. Diversity of structure and function of response regulator output domains. *Curr Opin Microbiol* **13**(2):150-159
51. **Ghuysen JM.** 1991. Serine beta-lactamases and penicillin-binding proteins. *Annu Rev Microbiol* **45**:37-67
52. **Gibson TJ.** 1984. Studies on the Epstein-Barr virus genome. PhD thesis Cambridge University, Cambridge, United Kingdom
53. **Gionchetti P, Rizzello F, Helwig U, Venturi A, Lammers KM, Brigidi P, Vitali B, Poggioli G, Miglioli M, Campieri M.** 2003. Prophylaxis of pouchitis onset with probiotic therapy: a double-blind, placebo-controlled trial. *Gastroenterology* **124**(5):1202-1209
54. **Girardin SE, Boneca IG, Viala J, Chamaillard M, Labigne A, Thomas G, Philpott DJ, Sansonetti PJ.** 2003. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* **278**(11):8869-8872
55. **Goffin C, Ghuysen JM.** 1998. Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. *Microbiol Mol Biol Rev* **62**(4):1079-1093
56. **Grangette C, Nutten S, Palumbo E, Morath S, Hermann C, Dewulf J, Pot B, Hartung T, Hols P, Mercenier A.** 2005. Enhanced antiinflammatory capacity of a *Lactobacillus*

- plantarum* mutant synthesizing modified teichoic acids. Proc Natl Acad Sci U S A **102**(29):10321-10326
57. **Grebe TW, Stock JB.** 1999. The histidine protein kinase superfamily. Adv Microb Physiol **41**:139-227
  58. **Grundling A, Schneewind O.** 2007. Synthesis of glycerol phosphate lipoteichoic acid in *Staphylococcus aureus*. Proc Natl Acad Sci U S A **104**(20):8478-8483
  59. **Gu J, Lu R, Liu X, Han W, Lei L, Gao Y, Zhao H, Li Y, Diao Y.** 2011. LysGH15B, the SH3b domain of staphylococcal phage endolysin LysGH15, retains high affinity to staphylococci. Curr Microbiol **63**(6):538-542
  60. **Hakenbeck R, Madhour A, Denapate D, Bruckner R.** 2009. Versatility of choline metabolism and choline-binding proteins in *Streptococcus pneumoniae* and commensal streptococci. FEMS Microbiol Rev **33**(3):572-586
  61. **Haldimann A, Fisher SL, Daniels LL, Walsh CT, Wanner BL.** 1997. Transcriptional regulation of the *Enterococcus faecium* BM4147 vancomycin resistance gene cluster by the VanS-VanR two-component regulatory system in *Escherichia coli* K-12. J Bacteriol **179**(18):5903-5913
  62. **Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J.** 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol **8**(2):R19
  63. **Henke JM, Bassler BL.** 2004. Three parallel quorum-sensing systems regulate gene expression in *Vibrio harveyi*. J Bacteriol **186**(20):6902-6914
  64. **Hill M, Deghmane AE, Segovia M, Zarantonelli ML, Tilly G, Blancou P, Bériou G, Josien R, Anegon I, Hong E, Ruckly C, Antignac A, El Ghachi M, Boneca IG, Taha MK, Cuturi MC.** 2011. Penicillin binding proteins as danger signals: meningococcal penicillin binding protein 2 activates dendritic cells through Toll-like receptor 4. PLoS One **6**(10):e23995
  65. **Hoch JA.** 2000. Two-component and phosphorelay signal transduction. Curr Opin Microbiol **3**(2):165-170
  66. **Hoch JA, Varughese KI.** 2001. Keeping signals straight in phosphorelay signal transduction. J Bacteriol **183**(17):4941-4949
  67. **Hooper LV, Gordon JI.** 2001. Commensal host-bacterial relationships in the gut. Science **292**(5519):1115-1118
  68. **Howell A, Dubrac S, Noone D, Varughese KI, Devine K.** 2006. Interactions between the YycFG and PhoPR two-component systems in *Bacillus subtilis*: the PhoR kinase phosphorylates the non-cognate YycF response regulator upon phosphate limitation. Mol Microbiol **59**(4):1199-1215
  69. **Huang YH, Ferrieres L, Clarke DJ.** 2006. The role of the Rcs phosphorelay in *Enterobacteriaceae*. Res Microbiol **157**(3):206-212
  70. **Hughes AH, Hancock IC, Baddiley J.** 1973. The function of teichoic acids in cation control in bacterial membranes. Biochem J **132**(1):83-93
  71. **Hynonen U, Palva A.** 2013. *Lactobacillus* surface layer proteins: structure, function and applications. Appl Microbiol Biotechnol **97**(12):5225-5243
  72. **Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenther D, Bovee D, Olson MV, Manoil C.** 2003. Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A **100**(24):14339-14344
  73. **Jayarathne P, Keenleyside WJ, MacLachlan PR, Dodgson C, Whitfield C.** 1993. Characterization of rcsB and rcsC from *Escherichia coli* O9:K30:H12 and examination of the role of the rcs regulatory system in expression of group I capsular polysaccharides. J Bacteriol **175**(17):5384-5394
  74. **Jenal U, Galperin MY.** 2009. Single domain response regulators: molecular switches with emerging roles in cell organization and dynamics. Curr Opin Microbiol **12**(2):152-160

75. **Jones AL, Needham RH, Clancy A, Knoll KM, Rubens CE.** 2003. Penicillin-binding proteins in *Streptococcus agalactiae*: a novel mechanism for evasion of immune clearance. *Mol Microbiol* **47**(1):247-256
76. **Kandler O, Weiss N.** 1986. Genus *Lactobacillus*. P H A Sheath, N S Maiz, M E Sharp, and J G Holt (ed), *Bergey's manual of systematic bacteriology* **2**:1209-1234
77. **Kankainen M, Paulin L, Tynkkynen S, von Ossowski I, Reunanen J, Partanen P, Satokari R, Vesterlund S, Hendrickx AP, Lebeer S, De Keersmaecker SC, Vanderleyden J, Hamalainen T, Laukkanen S, Salovuori N, Ritari J, Alatalo E, Korpela R, Mattila-Sandholm T, Lassig A, Hatakka K, Kinnunen KT, Karjalainen H, Saxelin M, Laakso K, Surakka A, Palva A, Salusjarvi T, Auvinen P, de Vos WM.** 2009. Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein. *Proc Natl Acad Sci U S A* **106**(40):17193-17198
78. **Kirby JR.** 2009. Chemotaxis-like regulatory systems: unique roles in diverse bacteria. *Annu Rev Microbiol* **63**:45-59
79. **Kleerebezem M, Boekhorst J, van Kranenburg R, Molenaar D, Kuipers OP, Leer R, Tarchini R, Peters SA, Sandbrink HM, Fiers MW, Stiekema W, Lankhorst RM, Bron PA, Hoffer SM, Groot MN, Kerkhoven R, de Vries M, Ursing B, de Vos WM, Siezen RJ.** 2003. Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc Natl Acad Sci U S A* **100**(4):1990-1995
80. **Kleerebezem M, Hols P, Bernard E, Rolain T, Zhou M, Siezen RJ, Bron PA.** 2010. The extracellular biology of the lactobacilli. *FEMS Microbiol Rev* **34**(2):199-230
81. **Kleerebezem M, Vaughan EE.** 2009. Probiotic and gut Lactobacilli and Bifidobacteria: molecular approaches to study diversity and activity. *Annu Rev Microbiol* **63**:269-290
82. **Koshland DE, Jr.** 1977. A response regulator model in a simple sensory system. *Science* **196**(4294):1055-1063
83. **Kovacikova G, Skorupski K.** 1999. A *Vibrio cholerae* LysR homolog, AphB, cooperates with AphA at the *tcpPH* promoter to activate expression of the ToxR virulence cascade. *J Bacteriol* **181**(14):4250-4256
84. **Kovacs-Simon A, Titball RW, Michell SL.** 2011. Lipoproteins of bacterial pathogens. *Infect Immun* **79**(2):548-561
85. **Krell T, Lacal J, Busch A, Silva-Jimenez H, Guazzaroni ME, Ramos JL.** 2010. Bacterial sensor kinases: diversity in the recognition of environmental signals. *Annu Rev Microbiol* **64**:539-559
86. **Kumar A, Sarkar SK, Ghosh D, Ghosh AS.** 2012. Deletion of penicillin-binding protein 1b impairs biofilm formation and motility in *Escherichia coli*. *Res Microbiol* **163**(4):254-257
87. **Kuo YC, Liu CF, Lin JF, Li AC, Lo TC, Lin TH.** 2013. Characterization of putative class II bacteriocins identified from a non-bacteriocin-producing strain *Lactobacillus casei* ATCC 334. *Appl Microbiol Biotechnol* **97**(1):237-246
88. **Lahooti M, Harwood CR.** 1999. Transcriptional analysis of the *Bacillus subtilis* teichuronic acid operon. *Microbiology* **145** ( Pt 12):3409-3417
89. **Lakowicz JR, Thompson RB.** 1983. Differential polarized phase fluorometric studies of phospholipid bilayers under high hydrostatic pressure. *Biochim Biophys Acta* **732**(2):359-371
90. **Landete JM, Garcia-Haro L, Blasco A, Manzanares P, Berbegal C, Monedero V, Zuniga M.** 2010. Requirement of the *Lactobacillus casei* MaeKR two-component system for L-malic acid utilization via a malic enzyme pathway. *Appl Environ Microbiol* **76**(1):84-95
91. **Laub MT, Goulian M.** 2007. Specificity in two-component signal transduction pathways. *Annu Rev Genet* **41**:121-145
92. **Layec S, Decaris B, Leblond-Bourget N.** 2008. Diversity of Firmicutes peptidoglycan hydrolases and specificities of those involved in daughter cell separation. *Res Microbiol* **159**(7-8):507-515
93. **Lazarevic V, Abellan FX, Moller SB, Karamata D, Mauel C.** 2002. Comparison of ribitol and glycerol teichoic acid genes in *Bacillus subtilis* W23 and 168: identical function, similar divergent organization, but different regulation. *Microbiology* **148**(Pt 3):815-824

94. **Lebeer S, Claes I, Tytgat HL, Verhoeven TL, Marien E, von Ossowski I, Reunanen J, Palva A, Vos WM, Keersmaecker SC, Vanderleyden J.** 2012. Functional analysis of *Lactobacillus rhamnosus* GG pili in relation to adhesion and immunomodulatory interactions with intestinal epithelial cells. *Appl Environ Microbiol* **78**(1):185-193
95. **Lebeer S, Claes IJ, Verhoeven TL, Vanderleyden J, De Keersmaecker SC.** 2011. Exopolysaccharides of *Lactobacillus rhamnosus* GG form a protective shield against innate immune factors in the intestine. *Microb Biotechnol* **4**(3):368-374
96. **Lebeer S, Vanderleyden J, De Keersmaecker SC.** 2008. Genes and molecules of lactobacilli supporting probiotic action. *Microbiol Mol Biol Rev* **72**(4):728-764, Table of Contents
97. **Letunic I, Doerks T, Bork P.** 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* **40**(Database issue):D302-305
98. **Livak KJ, Schmittgen TD.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**(4):402-408
99. **Lukat GS, McCleary WR, Stock AM, Stock JB.** 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc Natl Acad Sci U S A* **89**(2):718-722
100. **Maddocks SE, Oyston PC.** 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**(Pt 12):3609-3623
101. **Mandlik A, Swierczynski A, Das A, Ton-That H.** 2008. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol* **16**(1):33-40
102. **Marteyn B, West NP, Browning DF, Cole JA, Shaw JG, Palm F, Mounier J, Prevost MC, Sansonetti P, Tang CM.** 2010. Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature* **465**(7296):355-358
103. **Martinez-Hackert E, Stock AM.** 1997. Structural relationships in the OmpR family of winged-helix transcription factors. *J Mol Biol* **269**(3):301-312
104. **Mascher T.** 2006. Intramembrane-sensing histidine kinases: a new family of cell envelope stress sensors in Firmicutes bacteria. *FEMS Microbiol Lett* **264**(2):133-144
105. **Mascher T, Helmann JD, Uden G.** 2006. Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* **70**(4):910-938
106. **Mascher T, Zahner D, Merai M, Balmelle N, de Saizieu AB, Hakenbeck R.** 2003. The *Streptococcus pneumoniae* *cia* regulon: CiaR target sites and transcription profile analysis. *J Bacteriol* **185**(1):60-70
107. **Mazurkiewicz P, Tang CM, Boone C, Holden DW.** 2006. Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. *Nat Rev Genet* **7**(12):929-939
108. **McCracken VJ, Lorenz RG.** 2001. The gastrointestinal ecosystem: a precarious alliance among epithelium, immunity and microbiota. *Cell Microbiol* **3**(1):1-11
109. **Meberg BM, Paulson AL, Priyadarshini R, Young KD.** 2004. Endopeptidase penicillin-binding proteins 4 and 7 play auxiliary roles in determining uniform morphology of *Escherichia coli*. *J Bacteriol* **186**(24):8326-8336
110. **Molle V, Fujita M, Jensen ST, Eichenberger P, Gonzalez-Pastor JE, Liu JS, Losick R.** 2003. The Spo0A regulon of *Bacillus subtilis*. *Mol Microbiol* **50**(5):1683-1701
111. **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** ( Pt 10):2873-2883
112. **Mori H, Ito K.** 2001. The Sec protein-translocation pathway. *Trends Microbiol* **9**(10):494-500
113. **Navarre WW, Ton-That H, Faul KF, Schneewind O.** 1999. Multiple enzymatic activities of the murein hydrolase from staphylococcal phage phi11. Identification of a D-alanyl-glycine endopeptidase activity. *J Biol Chem* **274**(22):15847-15856
114. **Nelson DE, Young KD.** 2000. Penicillin binding protein 5 affects cell diameter, contour, and morphology of *Escherichia coli*. *J Bacteriol* **182**(6):1714-1721
115. **Neuhaus FC, Baddiley J.** 2003. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**(4):686-723

116. **Neville BA, Forde BM, Claesson MJ, Darby T, Coghlan A, Nally K, Ross RP, O'Toole PW.** 2012. Characterization of pro-inflammatory flagellin proteins produced by *Lactobacillus ruminis* and related motile Lactobacilli. *PLoS One* **7(7)**:e40592
117. **Nigro G, Rossi R, Commere PH, Jay P, Sansonetti PJ.** 2014. The cytosolic bacterial peptidoglycan sensor Nod2 affords stem cell protection and links microbes to gut epithelial regeneration. *Cell Host Microbe* **15(6)**:792-798
118. **Nikolskaya AN, Galperin MY.** 2002. A novel type of conserved DNA-binding domain in the transcriptional regulators of the AlgR/AgrA/LytR family. *Nucleic Acids Res* **30(11)**:2453-2459
119. **Palumbo E, Deghorain M, Cocconcelli PS, Kleerebezem M, Geyer A, Hartung T, Morath S, Hols P.** 2006. D-alanyl ester depletion of teichoic acids in *Lactobacillus plantarum* results in a major modification of lipoteichoic acid composition and cell wall perforations at the septum mediated by the Acm2 autolysin. *J Bacteriol* **188(10)**:3709-3715
120. **Peant B, LaPointe G, Gilbert C, Atlan D, Ward P, Roy D.** 2005. Comparative analysis of the exopolysaccharide biosynthesis gene clusters from four strains of *Lactobacillus rhamnosus*. *Microbiology* **151(Pt 6)**:1839-1851
121. **Perea Velez M, Verhoeven TL, Draing C, Von Aulock S, Pfitzenmaier M, Geyer A, Lambrichts I, Grangette C, Pot B, Vanderleyden J, De Keersmaecker SC.** 2007. Functional analysis of D-alanylation of lipoteichoic acid in the probiotic strain *Lactobacillus rhamnosus* GG. *Appl Environ Microbiol* **73(11)**:3595-3604
122. **Perraud AL, Weiss V, Gross R.** 1999. Signalling pathways in two-component phosphorelay systems. *Trends Microbiol* **7(3)**:115-120
123. **Pescaretti Mde L, Farizano JV, Morero R, Delgado MA.** 2013. A novel insight on signal transduction mechanism of RcsCDB system in *Salmonella enterica* serovar typhimurium. *PLoS One* **8(9)**:e72527
124. **Poyart C, Pellegrini E, Marceau M, Baptista M, Jaubert F, Lamy MC, Trieu-Cuot P.** 2003. Attenuated virulence of *Streptococcus agalactiae* deficient in D-alanyl-lipoteichoic acid is due to an increased susceptibility to defensins and phagocytic cells. *Mol Microbiol* **49(6)**:1615-1625
125. **Pragai Z, Allenby NE, O'Connor N, Dubrac S, Rapoport G, Msadek T, Harwood CR.** 2004. Transcriptional regulation of the *phoPR* operon in *Bacillus subtilis*. *J Bacteriol* **186(4)**:1182-1190
126. **Quach D, van Sorge NM, Kristian SA, Bryan JD, Shelver DW, Doran KS.** 2009. The CiaR response regulator in group B *Streptococcus* promotes intracellular survival and resistance to innate immune defenses. *J Bacteriol* **191(7)**:2023-2032
127. **Regulski K, Courtin P, Meyrand M, Claes IJ, Lebeer S, Vanderleyden J, Hols P, Guillot A, Chapot-Chartier MP.** 2012. Analysis of the peptidoglycan hydrolase complement of *Lactobacillus casei* and characterization of the major gamma-D-glutamyl-L-lysyl-endopeptidase. *PLoS One* **7(2)**:e32301
128. **Revilla-Guarinos A, Gebhard S, Alcantara C, Staron A, Mascher T, Zuniga M.** 2013. Characterization of a regulatory network of peptide antibiotic detoxification modules in *Lactobacillus casei* BL23. *Appl Environ Microbiol* **79(10)**:3160-3170
129. **Rhee S, Martin RG, Rosner JL, Davies DR.** 1998. A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc Natl Acad Sci U S A* **95(18)**:10413-10418
130. **Rokeach LA, Haselby JA, Hoch SO.** 1988. Molecular cloning of a cDNA encoding the human Sm-D autoantigen. *Proc Natl Acad Sci U S A* **85(13)**:4832-4836
131. **Russo TA, MacDonald U, Beanan JM, Olson R, MacDonald IJ, Sauberan SL, Luke NR, Schultz LW, Umland TC.** 2009. Penicillin-binding protein 7/8 contributes to the survival of *Acinetobacter baumannii* in vitro and in vivo. *J Infect Dis* **199(4)**:513-521
132. **Sambrook J, Fritsch EF, Maniatis T.** 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor, NY

133. **Sansonetti PJ, Hale TL, Dammin GJ, Kapfer C, Collins HH, Jr., Formal SB.** 1983. Alterations in the pathogenicity of *Escherichia coli* K-12 after transfer of plasmid and chromosomal genes from *Shigella flexneri*. *Infect Immun* **39**(3):1392-1402
134. **Sara M, Sleytr UB.** 2000. S-Layer proteins. *J Bacteriol* **182**(4):859-868
135. **Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P.** 2008. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol Rev* **32**(2):234-258
136. **Schachtsiek M, Hammes WP, Hertel C.** 2004. Characterization of *Lactobacillus coryniformis* DSM 20001T surface protein Cpf mediating coaggregation with and aggregation among pathogens. *Appl Environ Microbiol* **70**(12):7078-7085
137. **Schirner K, Marles-Wright J, Lewis RJ, Errington J.** 2009. Distinct and essential morphogenic functions for wall- and lipo-teichoic acids in *Bacillus subtilis*. *EMBO J* **28**(7):830-842
138. **Schneewind O, Model P, Fischetti VA.** 1992. Sorting of protein A to the staphylococcal cell wall. *Cell* **70**(2):267-281
139. **Schultz J, Milpetz F, Bork P, Ponting CP.** 1998. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**(11):5857-5864
140. **Sengupta R, Altermann E, Anderson RC, McNabb WC, Moughan PJ, Roy NC.** 2013. The role of cell surface architecture of lactobacilli in host-microbe interactions in the gastrointestinal tract. *Mediators Inflamm* **2013**:237921
141. **Seth A, Yan F, Polk DB, Rao RK.** 2008. Probiotics ameliorate the hydrogen peroxide-induced epithelial barrier disruption by a PKC- and MAP kinase-dependent mechanism. *Am J Physiol Gastrointest Liver Physiol* **294**(4):G1060-1069
142. **Shu CJ, Zhulin IB.** 2002. ANTA: an RNA-binding domain in transcription antitermination regulatory proteins. *Trends Biochem Sci* **27**(1):3-5
143. **Sidote DJ, Barbieri CM, Wu T, Stock AM.** 2008. Structure of the *Staphylococcus aureus* AgrA LytTR domain bound to DNA reveals a beta fold with an unusual mode of binding. *Structure* **16**(5):727-735
144. **Siezen R, Boekhorst J, Muscariello L, Molenaar D, Renckens B, Kleerebezem M.** 2006. *Lactobacillus plantarum* gene clusters encoding putative cell-surface protein complexes for carbohydrate utilization are conserved in specific gram-positive bacteria. *BMC Genomics* **7**:126
145. **Sinensky M.** 1974. Homeoviscous adaptation--a homeostatic process that regulates the viscosity of membrane lipids in *Escherichia coli*. *Proc Natl Acad Sci U S A* **71**(2):522-525
146. **Smit E, Oling F, Demel R, Martinez B, Pouwels PH.** 2001. The S-layer protein of *Lactobacillus acidophilus* ATCC 4356: identification and characterisation of domains responsible for S-protein assembly and cell wall binding. *J Mol Biol* **305**(2):245-257
147. **Sperandio B, Gautier C, McGovern S, Ehrlich DS, Renault P, Martin-Verstraete I, Guedon E.** 2007. Control of methionine synthesis and uptake by MetR and homocysteine in *Streptococcus mutans*. *J Bacteriol* **189**(19):7032-7044
148. **Steen A, Palumbo E, Deghorain M, Cocconcelli PS, Delcour J, Kuipers OP, Kok J, Buist G, Hols P.** 2005. Autolysis of *Lactococcus lactis* is increased upon D-alanine depletion of peptidoglycan and lipoteichoic acids. *J Bacteriol* **187**(1):114-124
149. **Stephenson K, Hoch JA.** 2002. Evolution of signalling in the sporulation phosphorelay. *Mol Microbiol* **46**(2):297-304
150. **Stock AM, Robinson VL, Goudreau PN.** 2000. Two-component signal transduction. *Annu Rev Biochem* **69**:183-215
151. **Strominger JL, Tipper DJ.** 1965. Bacterial cell wall synthesis and structure in relation to the mechanism of action of penicillins and other antibacterial agents. *Am J Med* **39**(5):708-721
152. **Sturme MH, Francke C, Siezen RJ, de Vos WM, Kleerebezem M.** 2007. Making sense of quorum sensing in lactobacilli: a special focus on *Lactobacillus plantarum* WCFS1. *Microbiology* **153**(Pt 12):3939-3947

153. **Takeda S, Fujisawa Y, Matsubara M, Aiba H, Mizuno T.** 2001. A novel feature of the multistep phosphorelay in *Escherichia coli*: a revised model of the RcsC --> YojN --> RcsB signalling pathway implicated in capsular synthesis and swarming behaviour. *Mol Microbiol* **40**(2):440-450
154. **Tallant T, Deb A, Kar N, Lupica J, de Veer MJ, DiDonato JA.** 2004. Flagellin acting via TLR5 is the major activator of key signaling pathways leading to NF-kappa B and proinflammatory gene program activation in intestinal epithelial cells. *BMC Microbiol* **4**:33
155. **Taylor BL, Zhulin IB.** 1999. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* **63**(2):479-506
156. **Thibessard A, Fernandez A, Gintz B, Leblond-Bourget N, Decaris B.** 2002. Effects of rodA and pbp2b disruption on cell morphology and oxidative stress response of *Streptococcus thermophilus* CNRZ368. *J Bacteriol* **184**(10):2821-2826
157. **Thiéry J-P.** 1967. Mise en évidence des polysaccharides sur coupes fines en microscopie électronique. *J Microscopie* **6**:987-1018
158. **Tiekink M, Kaditzky S, Valcheva R, Korakli M, Vogel RF, Ganzle MG.** 2005. Extracellular homopolysaccharides and oligosaccharides from intestinal lactobacilli. *J Appl Microbiol* **99**(3):692-702
159. **Tong X, Campbell JW, Balazsi G, Kay KA, Wanner BL, Gerdes SY, Oltvai ZN.** 2004. Genome-scale identification of conditionally essential genes in *E. coli* by DNA microarrays. *Biochem Biophys Res Commun* **322**(1):347-354
160. **Tremaroli V, Backhed F.** 2012. Functional interactions between the gut microbiota and host metabolism. *Nature* **489**(7415):242-249
161. **Trevors JT.** 2003. Fluorescent probes for bacterial cytoplasmic membrane research. *J Biochem Biophys Methods* **57**(2):87-103
162. **Ulrich LE, Koonin EV, Zhulin IB.** 2005. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol* **13**(2):52-56
163. **van Roosmalen ML, Geukens N, Jongbloed JD, Tjalsma H, Dubois JY, Bron S, van Dijk JM, Anne J.** 2004. Type I signal peptidases of Gram-positive bacteria. *Biochim Biophys Acta* **1694**(1-3):279-297
164. **Vollmer W, Blanot D, de Pedro MA.** 2008. Peptidoglycan structure and architecture. *FEMS Microbiol Rev* **32**(2):149-167
165. **von Ossowski I, Reunanen J, Satokari R, Vesterlund S, Kankainen M, Huhtinen H, Tynkkynen S, Salminen S, de Vos WM, Palva A.** 2010. Mucosal adhesion properties of the probiotic *Lactobacillus rhamnosus* GG SpaCBA and SpaFED pilin subunits. *Appl Environ Microbiol* **76**(7):2049-2057
166. **West AH, Stock AM.** 2001. Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci* **26**(6):369-376
167. **West NP, Sansonetti P, Mounier J, Exley RM, Parsot C, Guadagnini S, Prevost MC, Prochnicka-Chalufour A, Delepierre M, Tanguy M, Tang CM.** 2005. Optimization of virulence functions through glucosylation of *Shigella* LPS. *Science* **307**(5713):1313-1317
168. **Williams SB, Stewart V.** 1999. Functional similarities among two-component sensors and methyl-accepting chemotaxis proteins suggest a role for linker region amphipathic helices in transmembrane signal transduction. *Mol Microbiol* **33**(6):1093-1102
169. **Winkler ME, Hoch JA.** 2008. Essentiality, bypass, and targeting of the YycFG (VicRK) two-component regulatory system in gram-positive bacteria. *J Bacteriol* **190**(8):2645-2648
170. **Wolanin PM, Thomason PA, Stock JB.** 2002. Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol* **3**(10):REVIEWS3013
171. **Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, Sinha R, Gilroy E, Gupta K, Baldassano R, Nessel L, Li H, Bushman FD, Lewis JD.** 2011. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**(6052):105-108
172. **Yan F, Cao H, Cover TL, Whitehead R, Washington MK, Polk DB.** 2007. Soluble proteins produced by probiotic bacteria regulate intestinal epithelial cell survival and growth. *Gastroenterology* **132**(2):562-575

173. **Yang XF, Ji Y, Schneider BL, Reitzer L.** 2004. Phosphorylation-independent dimer-dimer interactions by the enhancer-binding activator NtrC of *Escherichia coli*: a third function for the C-terminal domain. *J Biol Chem* **279**(35):36708-36714
174. **Zahner D, Kaminski K, van der Linden M, Mascher T, Meral M, Hakenbeck R.** 2002. The ciaR/ciaH regulatory network of *Streptococcus pneumoniae*. *J Mol Microbiol Biotechnol* **4**(3):211-216
175. **Zhang YC, Zhang LW, Tuo YF, Guo CF, Yi HX, Li JY, Han X, Du M.** 2010. Inhibition of *Shigella sonnei* adherence to HT-29 cells by lactobacilli from Chinese fermented food and preliminary characterization of S-layer protein involvement. *Res Microbiol* **161**(8):667-672
176. **Zhou M, Boekhorst J, Francke C, Siezen RJ.** 2008. LocateP: genome-scale subcellular-location predictor for bacterial proteins. *BMC Bioinformatics* **9**:173
177. **Zuniga M, Gomez-Escoin CL, Gonzalez-Candelas F.** 2011. Evolutionary history of the OmpR/IIIa family of signal transduction two component systems in *Lactobacillaceae* and *Leuconostocaceae*. *BMC Evol Biol* **11**:34

# **Annexes**

Au cours de ma thèse, l'équipe a réalisé une collaboration scientifique avec l'équipe d'Aldo Corsetti et Giovana Suzzi (Université de Teramo, Italie), spécialistes de la flore microbienne des olives de tables et de la fermentation lactique correspondante qui est réalisée par la bactérie lactique *Lactobacillus pentosus*. L'objectif était d'identifier les gènes de cette bactérie impliqués dans sa capacité à croître dans ce milieu cumulant plusieurs facteurs de stress microbien (pH acide, pauvreté en nutriments, saumure NaCl, composé phénoliques antimicrobiens...). Une approche de génétique inverse globale, en utilisant le système P<sub>junc</sub>-TpaseIS<sub>1223</sub>, combinée à un système de criblage reposant sur un milieu combinant ces stress, a été réalisée et a permis d'identifier les mutants et donc les gènes nécessaires à la bactérie pour survivre dans ce milieu. Les mutants inaptes à la croissance sur ce milieu ont été caractérisés notamment dans leur capacité à former des biofilms à la surface des olives. Dans ce cadre, je me suis impliquée dans la réflexion méthodologique et dans la formation technique de l'étudiante Giorgia Perpetuini, qui a réalisé ce travail dans notre laboratoire. Les principaux résultats sont présentés dans cette annexe sous forme d'un article publié et d'un article en préparation.

Article 5. Identification of critical genes for growth in olive brine by transposon mutagenesis of *Lactobacillus pentosus* C11.

Perpetuini G, **Scornec H**, Tofalo R, Serrero P, Schirone M, Suzzi G, Corsetti A, Cavin J-F and Licandro-Seraut H. 2013. Appl. Environ. Microbiol. 79: 4568-4575.

Article 6. Ability of *Lactobacillus pentosus* C11 and the "olive brine adaptation" derived mutants to form biofilms on olive skin during fermentation.

Perpetuini G, Pham-Hoang BN, **Scornec H**, Tofalo R, Schirone M, Suzzi G, Corsetti A, Waché Y, Cavin J-F. and Licandro-Seraut H. Prochainement soumis.



## Identification of Critical Genes for Growth in Olive Brine by Transposon Mutagenesis of *Lactobacillus pentosus* C11

G. Perpetuini,<sup>a,b</sup> H. Scornec,<sup>b</sup> R. Tofalo,<sup>a</sup> P. Serror,<sup>d,e</sup> M. Schirone,<sup>a</sup> G. Suzzi,<sup>a</sup> A. Corsetti,<sup>a</sup> J. F. Cavin,<sup>b</sup> H. Licandro-Seraut<sup>b,c</sup>

Dipartimento Scienze degli Alimenti, Università di Teramo, Teramo, Italy<sup>a</sup>; UMR A 02102 PAM, AgroSup Dijon, and Université de Bourgogne, Dijon, France<sup>b</sup>; Unité de Pathogénie Microbienne Moléculaire, Institut Pasteur, Paris, France<sup>c</sup>; INRA, UMR1319 Micalis, Jouy-en-Josas, France<sup>d</sup>; AgroParisTech, UMR Micalis, Jouy-en-Josas, France<sup>e</sup>

Olive brine represents a stressful environment due to the high NaCl concentration, presence of phenolic compounds known as antimicrobials, and low availability of nutrients. Thus, only a few strains of lactic acid bacteria (LAB) are adapted to grow in and ferment table olives. To identify the mechanisms by which these few strains are able to grow in olive brine, *Lactobacillus pentosus* C11, a particularly resistant strain isolated from naturally fermented table olives, was mutagenized by random transposition using the P<sub>junC</sub>-TpaseI223 system (H. Licandro-Seraut, S. Brinster, M. van de Guchte, H. Scornec, E. Maguin, P. Sansonetti, J. F. Cavin, and P. Serror, *Appl. Environ. Microbiol.* 78:5417–5423, 2012). A library of 6,000 mutants was generated and screened for adaptation and subsequent growth in a medium, named BSM (brine screening medium), which presents the stressful conditions encountered in olive brine. Five transposition mutants impaired in growth on BSM were identified. Transposition occurred in two open reading frames and in three transcription terminators affecting stability of transcripts. Thus, several essential genes for adaptation and growth of *L. pentosus* C11 in olive brine were identified.

Table olives represent one of the most important fermented vegetables, as they are widespread both in the Mediterranean area and in Australia, the United States, and South America; they are highly appreciated for their sensory characteristics and nutritive value, and their use is extensive in all markets. Even if table olives represent an important economic source for the producing countries, the fermentation process is still empirical and craft based (1). Table olives can be elaborated in several ways, but the main industrial elaborations are the green Spanish style and the naturally black olives, or Greek style. These treatments aim to remove the natural bitterness of olives through the hydrolysis of oleuropein. In the Spanish style bitterness is removed by adding lye (NaOH solution), while in the Greek style olives are directly brined and the debittering is slow and only partial (1, 2). These processes are unpredictable and can lead to low-quality products. Many studies have focused on standardizing the quality of products through the use of starter cultures composed of lactic acid bacteria (LAB) and yeasts (3–8). Among yeasts, the species proposed as starters are *Candida diddensiae* (3), *Saccharomyces cerevisiae* (9), and *Debaryomyces hansenii* (10). Regarding LAB, *Lactobacillus plantarum* (11–13) and *Lactobacillus pentosus* (14, 15), which are very closely related species, are the main bacteria used as starters. They possess general stress resistance genes that have been identified in *L. plantarum* for other food fermentations (14). During the olive fermentation process, an ecological succession of different strains of these two species takes place. These strains have different biochemical and physiological characteristics and are adapted to the particular conditions of olive brines (15–17). In recent years, the interest of researchers was focused on the study of the antimicrobial effects of phenolic compounds present in table olives (18–20). Particular attention was paid to *L. plantarum*, while only few analyses were performed with *L. pentosus* (20). Some bacterial responses to phenolic compounds were deeply characterized, particularly the phenolic acid stress response (PASR) in *Pediococcus pentosaceus* (21), *L. plantarum* (22–24), and *Bacillus subtilis* (25, 26), which allows the bacteria to face ferulic, *p*-coumaric, and caffeic acid stresses.

The species of LAB encountered most frequently during table olive fermentation are *L. plantarum* and *L. pentosus* (17). Effectively, previous studies demonstrated that, at the end of fermentation of Itrana Bianca table olives (naturally fermented green table olives), the bacterial population was clearly dominated by *L. pentosus* and *L. plantarum* species (11–13). The mechanisms involved in the ability of these strains to grow in such a hostile environment are not known. So, the aim of this study was the identification of genes involved in the capability of *L. pentosus* C11 to adapt and to grow in olive brine and thus to perform olive fermentation. Therefore, by using random transposon mutagenesis, one of the most powerful approaches widely used for genetic characterization of bacterial phenotypes (25, 27, 28), a library of 6,000 random transposon mutants was generated using the P<sub>junC</sub>-TpaseI223 system (29) and screened for the ability to grow in olive brine, using BSM (brine screening medium).

### MATERIALS AND METHODS

**Strains, plasmids, and growth conditions.** Bacterial strains and plasmids used in this study are listed in Table 1. Multiplex PCR on the *recA* gene was performed to check the assignment of strain C11 to the *L. pentosus* species, as previously described (30). Wild-type (WT) *L. pentosus* C11 and its corresponding mutants were routinely grown either in liquid or in agar MRS medium (20 g/liter; Difco) at 37°C without shaking, whereas for several growth ability tests, liquid or agar (20 g/liter) YG medium composed of 10 g/liter yeast extract and 10 g/liter glucose was used as the basal medium. *Escherichia coli* cells were grown aerobically at 37°C in Luria-Bertani (LB) medium. Relevant antibiotic concentrations were as follows: 150 µg/ml erythromycin (Em) and 50 µg/ml ampicillin (Am) for *Esche-*

Received 10 April 2013 Accepted 14 May 2013

Published ahead of print 17 May 2013

Address correspondence to J. F. Cavin, cavinjf@u-bourgogne.fr, or H. Licandro-Seraut, Helene.Seraut@u-bourgogne.fr.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.01159-13

TABLE 1 Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
<b>Strains</b>		
<i>L. pentosus</i> C11	Isolated from table olives	University of Teramo culture collection
<i>E. coli</i> TG1	<i>supE hsdΔ5 thi Δ (lac-proAB) F'</i> [traD36 proAB <sup>+</sup> lacI <sup>q</sup> lacZΔM15]	54
<b>Plasmids</b>		
pV1129	Ap <sup>r</sup> Cm <sup>r</sup> , pV11056 containing P <sub>hibA</sub> -IS1223ΔIR	29
pV1110	Em <sup>r</sup> , pBR322ori, P <sub>junc</sub>	29

*richia coli*, and 10 μg/ml chloramphenicol (Cm) and 5 μg/ml Em for recombinant strains of *L. pentosus* C11 (Table 1). The olive-extracted phenolic compounds mix used for phenotypic analysis of selected mutants was obtained through high-performance liquid chromatography purification of olive oil vegetation waters (31) and contained 131 mg/g 3,4-dihydroxyphenylethanol (3,4-DHPEA), 20 mg/g *p*-hydroxyphenylethanol (*p*-HPEA), 64 mg/g verbascoside, 165 mg/g 3,4-DHPEA-EDA (dialdehydic form of elenolic acid linked with 3,4-dihydroxyphenylethanol), and 381 mg/g other phenolic compounds.

**Development of BSM.** A solid medium called brine screening medium was developed in this study to isolate mutants affected in their ability to grow in olive brine. In particular, the brine was taken from the end of fermentation of Itrana Bianca table olives (naturally fermented green table olives), when antimicrobial phenolic compounds reach the highest concentration. The main characteristics of this brine were previously described (2, 32). In particular, the brine (70 g/liter NaCl, pH 4.0) was mainly characterized by the presence of 3,4-DHPEA, *p*-HPEA, and verbascoside (32), and also glucose and lactic acid (2). In order to obtain a 100% concentration of brine in the agar medium, brine was 2-fold concentrated with a rotary Buchi vacuum apparatus at 52°C and then supplemented with 20 g/liter glucose and 20 g/liter yeast extract as nutrients. The pH was adjusted to 4.0 with HCl, corresponding to the general olive brine pH. Then, the supplemented brine was pasteurized at 65°C for 45 min, a treatment sufficient to destroy any vegetative microorganisms (determined by agar plating [data not shown]), and mixed with an equal volume of sterile melted agar at 40 g/liter in water just before pouring the plates. Mutants were spotted on BSM plates (diameter, 150 mm) by using a 96-solid-pin replicator and incubated for 72 h at 37°C. YG medium at pH 6.0 was used as the positive growth control. Absence of growth on BSM plates was confirmed in a replicate experiment.

**DNA techniques.** All DNA manipulations were performed according to standard procedures (33). Plasmids were isolated by using a Nucleospin plasmid miniprep kit (Euromedex). Ligation and restriction analysis were carried out according to the manufacturer's instructions. PCR was performed using 0.1 unit of Platinum *Taq* DNA polymerase high fidelity (Invitrogen), according to the manufacturer's recommendations, in an automatic thermocycler (Bio-Rad).

**Transposon mutagenesis.** *L. pentosus* C11 and *E. coli* were transformed by electroporation as described by Aukrust et al. (34) and Dower et al. (35), respectively, using a GenePulser and a pulse controller apparatus (Bio-Rad). Plasmids were prepared from *E. coli* by using a Nucleospin plasmid miniprep kit (Euromedex). Mutagenesis was performed using the P<sub>junc</sub>-TpaSelS1223 system as previously described (29) with the following modifications. Electrocompetent cells of *L. pentosus* C11 were first electroporated with pV1129, which provided the transposase of IS1223, plated onto MRS agar with Cm at 10 μg/ml, and incubated at 37°C for 48 h. Electrocompetent cells of a pV1129 transformant were then transformed with pV1110. The cells were plated on MRS agar plates supplemented with 5 μg/ml Em and incubated at 42°C for 48 h to select integrants.

TABLE 2 Primers

Use and name <sup>a</sup>	Sequence (5'→3')
<b>pV1110 target sequencing</b>	
IRR6	TCACCGTCATCACCGAAACG
IRL6	GCCGACTAGTGATTAATAAC
<b>qRT-PCR</b>	
<i>obaA</i> F	GATTGGGCTCGTCCTAGTCA
<i>obaA</i> R	TATTGTTGGAAGCCGTCGAT
<i>obaB</i> F	CTTGACCGTTGTCACCCAAT
<i>obaB</i> R	AGAAACAGTGCGGGTTCAAA
<i>gpi</i> F	ATCGGGATTGGTGGTTCATA
<i>gpi</i> R	TGTGGGAATTTACGGTCTTCA
<i>obaC</i> F	CGTTTAAACAACGCTGAGCAA
<i>obaC</i> R	AACCCTTACCACAACAAGC
<i>obaE</i> F	ATAGCGACAGCACCTGCAC
<i>obaE</i> R	CGATGATTTGGTACCGGAAC
<i>obaF</i> F	TCAGGCACCATAAGCATCG
<i>obaF</i> R	TGCAACCGAATTAACAGGA
<i>enoA1</i> F	GCAAGGTAGTCCGTGTTCCGT
<i>enoA1</i> R	ACTGGGAAGACTGGCAAATG

<sup>a</sup> F, forward; R, reverse.

**Sequence analysis and mapping of transposon insertion site.** Genomic DNA was digested simultaneously with ClaI and BstBI (Fermentas), and ligation was performed using T4 DNA ligase HC (Fermentas) according to the manufacturer's instructions. The resulting ligation products were directly transformed into the *E. coli* TG1 strain, in which circularized fragments that contain the transposon replicate as plasmids. Plasmids were isolated from selected transformants and subjected to sequencing reactions (GATC Biotech) using the primers IRR6 and IRL6, which target the transposon sequence extremities (Table 2). Identification of transposon target sequences was performed with the BLAST software from the National Center for Biotechnology Information (<http://blast.ncbi.nlm.nih.gov/>).

**RNA extraction and qRT-PCR analysis.** Total RNA from 10 ml of culture was extracted after bead beating disruption using the Tri reagent method (Sigma) as previously described (24). Quantitative reverse transcriptase PCR (qRT-PCR) and calculations of relative transcript levels (RTLs) were carried out as previously described elsewhere (24), using the primers listed in Table 3. The genes *tpiA* and *rpoD* were used as internal calibrators for all qRT-PCR analyses.

## RESULTS

**Transposon mutagenesis of *L. pentosus* C11.** *L. pentosus* C11, isolated from olive brine (unpublished data), was mutagenized

TABLE 3 Analysis of transposon integration in 10 randomly selected *L. pentosus* C11 mutants based on sequencing of the transposon target and BLAST analysis with the *L. pentosus* IG1 genome sequence

Mutant no.	Locus of pV1110 insertion
1	IGR <i>lpent_00526/lpent_00527</i>
2	ORF <i>lpent_00833</i>
3	ORF <i>lpent_00835</i>
4	IGR <i>lpent_01101/lpent_01102</i>
5	IGR <i>lpent_01190/lpent_01192</i>
6	IGR <i>lpent_01719/lpent_01720</i>
7	IGR <i>lpent_01772/lpent_01773</i>
8	ORF <i>lpent_02891</i>
9	IGR <i>lpent_02977/lpent_02978</i>
10	IGR <i>lpent_03116/lpent_03117</i>

Perpetuini et al.

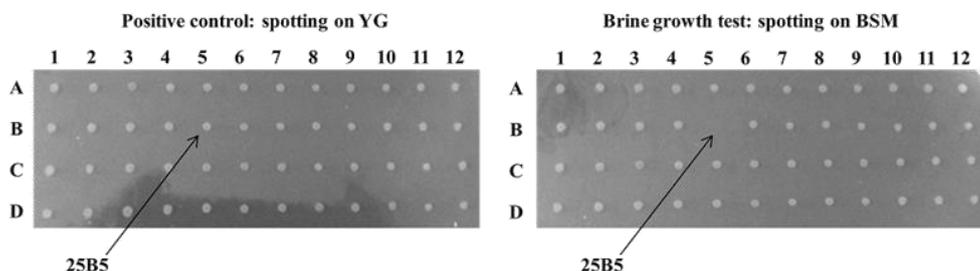


FIG 1 Photographs of plates obtained during screening of 48 *L. pentosus* C11 mutants for their sensitivity to olive brine after replication from fresh titration microplate individual cultures on either YG or BSM and incubation at 37°C for 72 h.

using the  $P_{\text{junc}}\text{-TpaseIS1223}$  system, which was specially designed for LAB random transposon mutagenesis and successfully used for *Lactobacillus casei* mutagenesis (29). This system is composed of the thermosensitive pVI129 plasmid, which is used for transient expression of the IS1223 transposase and the nonreplicative plasmid pVI110, which contains dedicated inverted repeat sequences that require the transposase for pVI110 random integration into the genome. The segregational stability of pVI129 in *L. pentosus* at 42°C was estimated to be 80% per generation by using the calculation method of Heap et al. (36). Thus, the mutagenesis procedure applied for *L. pentosus* was the same as that described for *L. casei*, but the final incubation was at 42°C. With approximately 4,000 mutants obtained from  $10^9$  viable cells (carrying pVI129) transformed with 1  $\mu\text{g}$  of pVI110, the transposition efficiency was similar to that of *L. casei* (29). Less than 10  $\text{Em}^r$  colonies were obtained for WT *L. pentosus* electroporated with pVI110, confirming the necessity for pVI129 to promote pVI110 integration. Eighteen pVI110 integration mutants were randomly selected and analyzed by Southern blotting with a pVI110-specific probe as previously performed (29). Southern blot analysis indicated that pVI110 was integrated randomly at a single locus of genomic DNA of *L. pentosus* C11 (data not shown). This was confirmed by sequencing of the genomic DNA target of pVI110 in 10 randomly selected mutants (Table 3). The transposon was integrated into different loci, strongly supporting the randomness of pVI110 integration in *L. pentosus*. Thirty percent of the mutants were disrupted in open reading frames (ORFs) (Table 3). Moreover, the target sequence of pVI110 was reanalyzed for 5 of the 10 randomly selected mutants after 50 generations and gave the same results, demonstrating the stability of pVI110 integration. Altogether, these results confirmed the efficiency and the randomness of the transposon mutagenesis in *L. pentosus* when using the  $P_{\text{junc}}\text{-TpaseIS1223}$  system. Thus, a collection of 6,000 colonies, in which about one-third of the mutants in an ORF are expected, were randomly picked and prepared in 96-well plates for phenotypic screening.

**Screening for mutants unable to grow on olive BSM.** In order to select mutants from the mutagenized library of *L. pentosus* C11 affected in their capacity to grow in brine, a solid agar medium named brine screening medium was developed by supplementing YG-agar medium with olive brine (see Materials and Methods). Mutants of the library were spotted on BSM plates, with a positive growth control on YG plates (Fig. 1). After 48 h of incubation at 37°C, of the 6,000 mutants screened, 5 failed to grow on BSM but grew on YG plates. This inability was also confirmed in BSM broth (BSM without agar): there was no increase of the optical density at

600 nm ( $\text{OD}_{600}$ ) monitored during the 48 h following the inoculation at an  $\text{OD}_{600}$  of 0.1.

#### Phenotypic analysis of the selected brine-sensitive mutants.

Phenolic compounds and osmotic stress triggered by salt concentration, which can reach 70 g/liter, are the main stresses encountered in olive brine. To establish the main factors responsible for mutant growth inhibition in brine and to better understand the niche adaptation phenomenon, the growth of the 5 selected mutants was compared to that of WT *L. pentosus* under the main stress conditions encountered in olive brine. Bacterial growth was tested in YG broth (pH 4) containing an increasing concentration of NaCl and supplemented or not with either oleuropein (0.2 g/liter) or an olive-extracted phenolic compounds mix (0.4 g/liter). Only a few differences between the growth of WT *L. pentosus* and mutants were detected in YG broth at pH 4 after 24 h, in the absence of stress conditions (Fig. 2) or with 10 g/liter NaCl (data not shown), with a similar final  $\text{OD}_{600}$  (about 1.3) for WT *L. pentosus* and mutants. These results demonstrated that the identified mutants were not sensitive to acidic conditions nor to 10-g/liter NaCl supplementation. The growth of all mutants was partially inhibited with 20 g/liter NaCl (Fig. 2), and none of the 5 mutants was able to grow in the presence of 30 g/liter NaCl or higher concentrations, contrary to findings with WT *L. pentosus* (data not shown).

Contrary to WT *L. pentosus*, the growth of the 5 brine-sensitive mutants was affected in the presence of the olive-extracted phenolic compound mix (Fig. 2). This result agrees with that obtained for *L. pentosus* ATCC 8041, which was able to grow in the presence of different phenolic compounds, such as hydroxytyrosol and its glucosides, oleoside, tyrosol, secoxyloganin, secologanoside, and oleuropein (20). Supplementation with 20 g/liter NaCl increased the inhibition triggered by phenolic compounds for all mutants (final  $\text{OD}_{600}$  of about 0.3), whereas it had no consequence on WT *L. pentosus* growth (Fig. 2). These observations confirmed that the inability of the mutants to grow in brine results from the combined stress induced by phenolic compounds and NaCl, as reported for other LAB strains not well adapted to this inhospitable environment (37). For example, oleuropein, one of the most abundant phenolic compounds in olive pulp, exerts an inhibitory effect on bacterial growth by a mechanism that remains unknown (20). It is suspected to induce leakage of glutamate and inorganic phosphate from bacterial cells, cause degradation of the cell wall itself (19, 38), and also cause a decrease in the ATP content of cells (39). The growth of the five brine-sensitive mutants was inhibited at the same level, with a final  $\text{OD}_{600}$  of 0.6 in YG containing 0.2 g/liter oleuropein, with or without 20 g/liter NaCl (Fig. 2). Notice-

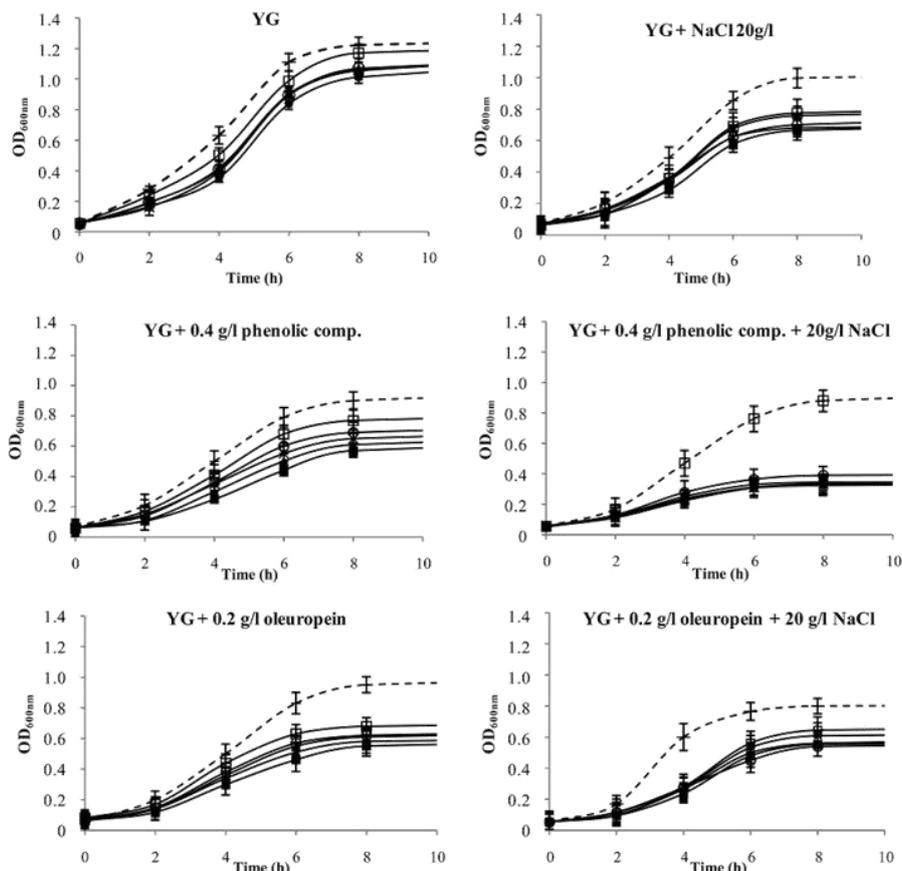


FIG 2 Growth of WT *L. pentosus* C11 (dotted line connected plus signs) and brine-sensitive mutants 25B5 (■), 20B10 (▲), 31B11 (×), 51D12 (○), and 42G1 (□) under different stress conditions with phenolic compounds, oleuropein, and NaCl.

ably, the inhibition of the mutant was particularly strong during the first hours of growth, in comparison with WT *L. pentosus* (Fig. 2).

**From phenotype to olive brine adaptation genes.** Transposon chromosomal targets were determined by sequencing and comparison of the insertion site sequence with *L. pentosus* IG1 genomic sequence using BLAST database searches (NCBI). The disrupted genes were named *obaA* to *obaE*, for olive brine adaptation, when no putative function was already attributed.

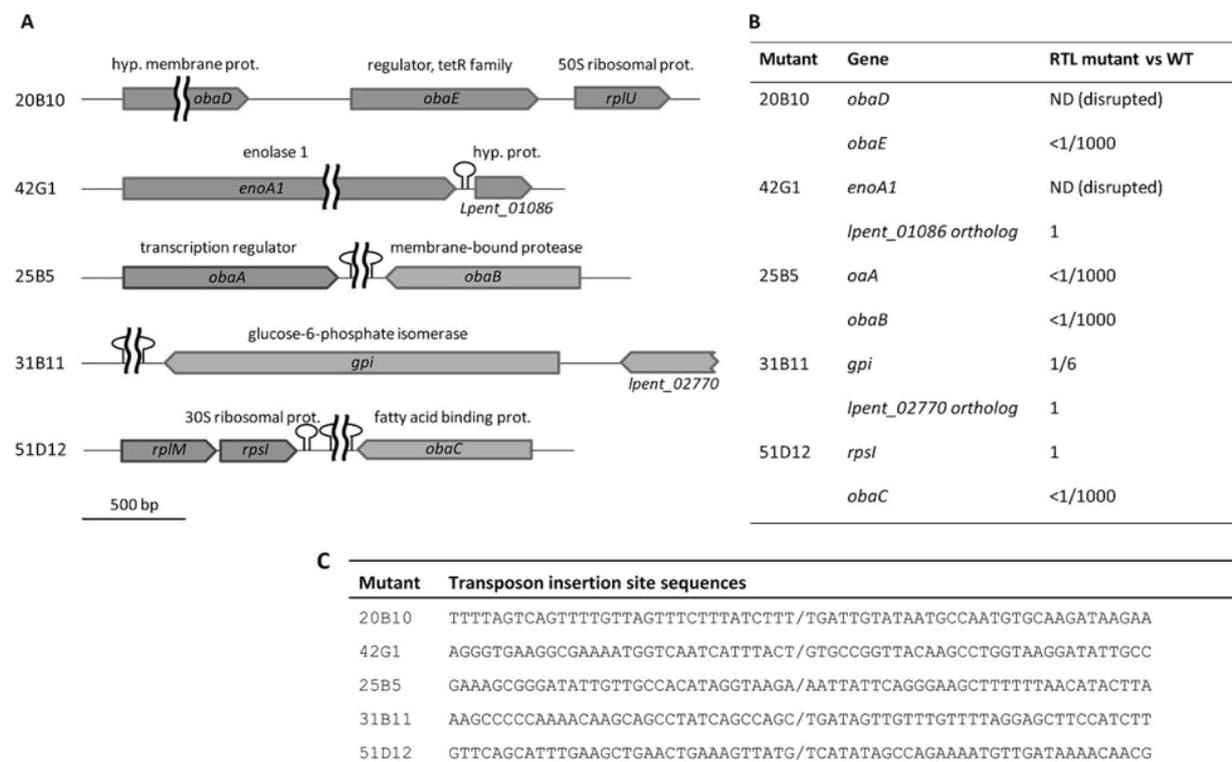
In mutant 20B10, the transposon was inserted in *obaD*, an ortholog of *lpent\_00392*, which encodes a small hypothetical integral membrane protein of unknown function. The *obaD* gene seems to be specific to *L. pentosus/L. plantarum* species, since neither ortholog nor even homologs of the ObaD-encoded protein were found in other bacterial species. This gene is located 379 bp upstream of *obaE*, an ortholog of *lpent\_00391*, which encodes a transcriptional regulator of the Tet<sup>r</sup> family. Despite the quite large intergenic region (IGR) between *obaD* and *obaE*, no putative transcriptional terminator was found in this IGR, and RT-PCR with two primers designed for *obaD* and *obaE*, respectively, on WT *L. pentosus* also indicated that *obaD* was cotranscribed with *obaE* (data not shown). The consequence of transposon integration on *obaE* was evaluated by qRT-PCR performed on mutant cultures

grown in MRS until the early stationary phase of growth, using WT *L. pentosus* under the same culture condition as the reference (Fig. 3). In the 20B10 mutant, *obaE* transcripts were under the detection level (lower than 1/1,000 of that measured for the WT), indicating that *obaE*, in addition to *obaD*, is silenced in this mutant.

The mutant 42G1 is disrupted for *enoA1*, an ortholog of *lpent\_01085*, which is predicted to encode an enolase (EnoA1), an essential enzyme for glycolysis (Fig. 3). *L. pentosus* C11 is likely to carry another enolase gene, as is the case for the sequenced *L. plantarum* strains, which would explain the fact that this disruption is not lethal. The RTL of the gene downstream of *enoA1*, the *lpent\_01086* ortholog, was not modified by the transposon integration in *enoA* (RTL of 1), and a putative transcription terminator was found downstream of *enoA1* (Fig. 3).

In mutants 31B11, 51D12, and 25B5, the transposon was integrated into the predicted transcription terminators (TT) (Fig. 3). In mutant 31B11, the transposon was integrated into the TT of *gpi*, an ortholog of *lpent\_02771*, which encodes glucose-6-phosphate isomerase, whereas in mutant 51D12 it is integrated into the TT of *obaC*, an ortholog of *lpent\_00851*, which encodes a putative fatty acid binding protein. The transposon of mutant 25B5 was integrated into a TT that could serve for the two convergent genes

Perpetuini et al.



**FIG 3** (A) Schematic representation of transposon integrations in the 5 olive brine-sensitive mutants. Transposon integration is represented by a vertical, curved double line. (B) RTLs of genes of the locus affected by the transposon, using WT *L. pentosus* C11 as the reference. ND, not determined (gene disrupted). (C) Transposon insertion site sequences.

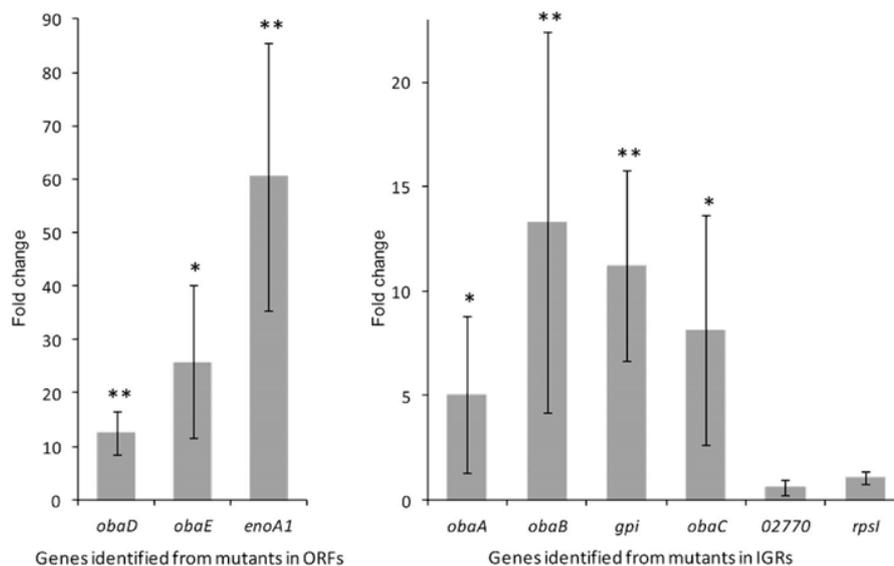
*obaA*, an ortholog of *lpent\_01150*, which encodes a putative redox-sensitive transcription regulator, and *obaB*, an ortholog of *lpent\_01149*, which encodes a putative membrane-bound protease.

Olive brine sensitivity of IGR integrants could be due to polar effects of the transposon on the flanking genes. To check this hypothesis, qRT-PCR analysis of genes flanking the transposon was performed for each mutant, using WT *L. pentosus* as the reference (Fig. 3). The RTL of *gpi* in mutant 31B11 was 6-fold lower than in the WT, while transposon insertion did not change the RTL of the upstream gene, the *lpent\_02770* ortholog. The lower *gpi* RTL could be explained by an alteration of the transcript stability, as bacterial single-stranded RNA is rapidly degraded from the 3' end (40, 55). The consequence should be a decrease of the glucose-6-phosphate isomerase (GPI) concentration in the cell, which could alter carbohydrate-related metabolic functions. As Gpi plays a central role in carbohydrate metabolism, a stronger reduction of the *gpi* RTL, or its inactivation by transposon insertion, would have been lethal. Interestingly, these results illustrate an unexpected aspect of IGR mutants, in which in some cases transposon integration into a TT reduces the transcript level of an essential gene without affecting growth and viability under nonstressful conditions. The transcript level of *obaC* was under the detection level in mutant 51D12 (<1/1,000 of that in WT), while the RTL of the divergently transcribed gene, *rpsI*, did not change (Fig. 3). This is consistent with the presence in the IGR for these two genes of two independent hairpin structures predicted to be transcription

terminators. The drastic reduction of the *obaC* RTL could be due to instability of *obaC* mRNA presenting an altered TT. Furthermore, a hairpin structure downstream of *obaC* displays high similarity level with the RF01676/P31 noncoding RNA family (<http://rfam.sanger.ac.uk/>). Interestingly, in mutant 25B5, transposon insertion reduced drastically the RTL of the two flanking genes, *obaA* and *obaB*. No alignment with noncoding RNA of a known family was found, but their presence cannot be excluded. From a functional point of view, 51D12 can be considered an *obaC* knockout mutant and 25B5 as a double *obaA-obaB* knockout mutant.

Despite the fact that we did not obtain two or more mutants in each of the five genes identified in this screening, indicating that testing of 6,000 mutants is not sufficient to reach saturation, our drastic screening conditions allowed us to identify some of the most essential genes for growth in olive brine.

Transcriptional analysis of the genes interrupted, or silenced by transposon integration, was carried out on WT *L. pentosus* after 16 h of growth in BSM broth, or in YG broth for the reference condition (Fig. 4). All genes identified in olive brine-sensitive mutants displayed a significantly higher RTL in BSM broth than in YG broth, with particularly high increases for *enoA1* (60-fold) and *obaE* (35-fold) (Fig. 4). This analysis reinforces their role in the resistance to the olive brine stress and in olive fermentation adaptability and demonstrates that this adaptation results in an upregulation of *oba* gene transcription in WT *L. pentosus*.



**FIG 4** Relative transcript levels of *L. pentosus* C11 genes after 16 h of growth in BSM broth. Transcript levels of each gene are expressed as the relative fold change, with YG medium as the reference condition (fold change = 1). Four biological repeats were performed, and bars indicate standard deviations. Statistical analysis was performed using the unpaired Student *t* test: \*,  $P < 0.05$ ; \*\*,  $P < 0.005$ .

## DISCUSSION

Olive brine represents a stressful environment due to the high NaCl concentration, presence of phenolic compounds considered antimicrobials, and low availability of nutrients. It was observed that the inhibitory effect of olive phenolic compounds on LAB growth was higher when they were associated with NaCl, showing a combined effect in the inhibition (37).

Oleuropein and hydroxytyrosol (the product of oleuropein hydrolysis) are the most abundant phenolic compounds found in olives, and they show antimicrobial effects (20). The mechanisms of bacterial growth inhibition have not been elucidated, and often controversial results have been obtained, depending on the different antimicrobial assays used (19). However, several studies agreed that oleuropein and its hydrolysis products could impact cell wall structure. These compounds could induce leakage of glutamate and inorganic phosphate from bacterial cells, as well as the degradation of the cell wall itself (19, 38) and changes in the typical bacillary structure of Gram-positive bacteria (38). Moreover, oleuropein has been reported to cause a decrease in the ATP content of the cells without affecting the rate of glycolysis (39).

The response to osmotic stress implies involvement of several genes whose expression, coordinated by osmosensing regulation, leads to an adjustment of cytoplasmic properties, cell turgor, hydration, and thus protein activity (42, 43). The sequence analysis of *L. pentosus* IG1, a strain isolated from Spanish-style green olive fermentations, revealed that this strain presents 16 putative two-component regulatory systems, which may reflect an extensive ability to adapt to changing environmental conditions (56).

The stress response in LAB and particularly in *Lactobacillus* has been reviewed recently (44). The genes involved in the generalized stress response in Gram-positive bacteria and LAB appear to be highly conserved (45). Thus, the ability of only some strains of *L. plantarum* and *L. pentosus* to face olive brine stresses is probably related to the presence of a genotypic diversity in the “lifestyle

adaptation island,” which is a chromosomal region suggested to be involved in niche adaptation (45).

The strategy used in this work allowed us to select 5 mutants of *L. pentosus* C11 unable to grow in olive brine because of the simultaneous presence of inhibitory concentrations of NaCl and the phenolic compounds that characterize this medium.

Thanks to genetic and transcriptomic analyses of mutants, *enoA1*, *gpi*, and *obaC* were clearly identified as essential genes for growth in olive brine. Concerning mutants 20B10 and 25B5, our results did not discriminate between *obaD* versus *obaE* or *obaA* versus *obaB*, respectively, since these mutants can be considered double knockouts and since these four genes are upregulated in the presence of olive brine. Further studies are required to establish which one of *obaA* or *obaB* is needed for adaptation to olive brine.

Gene expression analysis demonstrated that all *oba* genes are upregulated in the presence of olive brine, reinforcing the evidence of their involvement in brine stress resistance and growth in this medium. These genes can be classified into three groups: the first group includes *gpi* and *enoA1*, both of which encode enzymes involved in energetic metabolism; the second group includes *obaB*, *obaC*, and *obaD*, which encode putative membrane proteins; the third group includes *obaA* and *obaE*, which encode proteins with a putative regulatory function.

Under stress conditions, carbon metabolism, which is involved in energy production and saving, can be disturbed, as reported for *L. plantarum* challenged with tannic acids (46) and for *Lactobacillus rhamnosus* in the presence of acid stress (47). The enolase gene is generally one of the upregulated metabolic genes in response to stress; for instance, its expression is induced with decreasing temperatures of growth in *Lactobacillus helveticus* (48) and in a low-pH environment for *Lactobacillus reuteri* (49). So, enolase gene disruption, as well as a glucose-6-P-isomerase (*gpi*) RTL decrease, could affect the efficiency of metabolic pathways involved

in stress adaptation. Amazingly, enolase 1 has also been described as a surface protein able to bind fibronectin in *L. plantarum* (50), indicating that other roles for enolase 1 in sensing the environment or in interacting with olive brine molecules are possible.

The *obaD/obaE* operon encodes two putative proteins of unknown function. ObaD, a putative membrane protein without an identified conserved domain, is exclusively found in *L. plantarum* and *L. pentosus* species, while ObaE is a putative transcription regulator of the Tet<sup>r</sup> family. Thus, ObaD/ObaE could constitute an element of a system for sensing and responding to environmental changes like osmotic pressure variations and appearance of phenolic compounds. This is consistent with previous work that indicated that phenolic compounds are responsible for alterations of the cell wall and modifications of fatty acid membrane composition in *L. plantarum* (51).

The *L. plantarum* ortholog of *obaB*, which encodes a membrane-bound protease CAAX family protein, is upregulated in the presence of ethanol (52). It has been proposed that this protein could be involved in the lipid ordering and bilayer stability and it could influence permeability, fluidity, and the functioning of membrane-embedded enzymes (52). Regarding *obaC*, which encodes a putative DegV family protein, its cellular function remains unknown, but it could be involved in lipid transport or in fatty acid metabolism (53).

In this study, five mutants unable to grow in olive brine were isolated, and their analysis led to the identification of at least 5 genes essential for the growth of *L. pentosus* under olive fermentation conditions, which are characterized by a particularly high concentration of salt and phenolic compounds. The identified genes encode proteins involved in carbohydrate metabolism, membrane structure and function, or regulation of gene expression. As *obaD* encodes a putative membrane protein strictly specific to *L. pentosus/plantarum* species, it is one of the key elements that deserves further characterization in order to understand the particularly efficient adaptation mechanisms in these two closely related species to environments that present these specific stresses, which are found in many fermented food processes and also natural ecosystems.

#### ACKNOWLEDGMENTS

This work was part of the project 580 “Tracciabilità, certificazione e tutela della qualità dell’olio di oliva e delle olive da tavola—Azione 4d” supported by grant 582 from UNAPROL (867/2008 Misura 4). G. Perpetuini was the beneficiary of a grant financed by the European Social Fund (FSE). This work was also supported by the Scientific Council of AgroSup Dijon and Conseil Régional de Bourgogne.

#### REFERENCES

- Aponte M, Ventorino V, Blaiotta G, Volpe G, Farina V, Avellone G, Lanza CM, Moschetti G. 2010. Study of green Sicilian table olive fermentations through microbiological, chemical and sensory analyses. *Food Microbiol.* 27:162–170.
- Tofalo R, Schirone M, Perpetuini G, Angelozzi G, Suzzi G, Corsetti A. 2012. Microbiological and chemical profiles of naturally fermented table olives and brines from different Italian cultivars. *Antonie Van Leeuwenhoek* 102:121–131.
- Hurtado A, Reguant C, Bordons A, Rozes N. 2010. Evaluation of a single and combined inoculation of a *Lactobacillus pentosus* starter for processing cv. *Arbequina* natural green olives. *Food Microbiol.* 27:731–740.
- De Bellis P, Valerio F, Sisto A, Lonigro SL, Lavermicocca P. 2010. Probiotic table olives: microbial populations adhering on olive surface in fermentation sets inoculated with the probiotic strain *Lactobacillus paracasei* IMPC2.1 in an industrial plant. *Int. J. Food Microbiol.* 140:6–13.
- Panagou EZ, Schillinger U, Franz CM, Nychas GJ. 2008. Microbiological and biochemical profile of cv. *Conservolea* naturally black olives during controlled fermentation with selected strains of lactic acid bacteria. *Food Microbiol.* 25:348–358.
- Perricone M, Bevilacqua A, Corbo MR, Sinigaglia M. 2010. Use of *Lactobacillus plantarum* and glucose to control the fermentation of “Bella di Cerignola” table olives, a traditional variety of Apulian region (southern Italy). *J. Food Sci.* 75:M430–M436.
- Corsetti A, Perpetuini G, Schirone M, Tofalo R, Suzzi G. 2012. Application of starter cultures to table olive fermentation: an overview on the experimental studies. *Front. Microbiol.* 3:248.
- Botta C, Coccolin L. 2012. Microbial dynamics and biodiversity in table olive fermentation: culture-dependent and -independent approaches. *Front. Microbiol.* 3:245.
- Segovia Bravo KA, Arroyo Lopez FN, Garcia Garcia P, Duran Quintana MC, Garrido Fernandez A. 2007. Treatment of green table olive solutions with ozone. Effect on their polyphenol content and on *Lactobacillus pentosus* and *Saccharomyces cerevisiae* growth. *Int. J. Food Microbiol.* 114:60–68.
- Tsapatsaris S, Kotzekidou P. 2004. Application of central composite design and response surface methodology to the fermentation of olive juice by *Lactobacillus plantarum* and *Debaryomyces hansenii*. *Int. J. Food Microbiol.* 95:157–168.
- Leal-Sanchez MV, Ruiz-Barba JL, Sanchez AH, Rejano L, Jimenez-Diaz R, Garrido A. 2003. Fermentation profile and optimization of green olive fermentation using *Lactobacillus plantarum* LPCO10 as a starter culture. *Food Microbiol.* 20:421–430.
- Lamzira Z, Asehraou A, Brito D, Oliveira M, Faid M, Peres C. 2005. Reducing the bloater spoilage during lactic fermentation of Moroccan green olives. *Food Technol. Biotechnol.* 43:373–377.
- Marsilio V, Seghetti L, Iannucci E, Russi F, Lanza B, Felicioni M. 2005. Use of a lactic acid bacteria starter culture during green olive (*Olea europaea* L. cv. *Ascolana tenera*) processing. *J. Sci. Food Agric.* 85:1084–1090.
- Bron PA, Wels M, Bongers RS, van Bokhorst-van de Veen H, Wiersma A, Overmars L, Marco ML, Kleerebezem M. 2012. Transcriptomes reveal genetic signatures underlying physiological variations imposed by different fermentation conditions in *Lactobacillus plantarum*. *PLoS One* 7:e38720. doi:10.1371/journal.pone.0038720.
- Ruiz-Barba JL, Cathcart DP, Warner PJ, Jimenez-Diaz R. 1994. Use of *Lactobacillus plantarum* LPCO10, a bacteriocin producer, as a starter culture in Spanish-style green olive fermentations. *Appl. Environ. Microbiol.* 60:2059–2064.
- Ruiz-Barba JL, Jimenez-Diaz R. 2012. A novel *Lactobacillus pentosus*-paired starter culture for Spanish-style green olive fermentation. *Food Microbiol.* 30:253–259.
- Hurtado A, Reguant C, Bordons A, Rozes N. 2012. Lactic acid bacteria from fermented table olives. *Food Microbiol.* 31:1–8.
- Rodriguez H, Curiel JA, Landete JM, de las Rivas B, Lopez de Felipe F, Gomez-Cordoves C, Mancheno JM, Munoz R. 2009. Food phenolics and lactic acid bacteria. *Int. J. Food Microbiol.* 132:79–90.
- Landete MJ, Curiel AJ, Rodriguez H, Rivas BDL, Munoz R. 2008. Study of the inhibitory activity of phenolic compounds found in olive products and their degradation by *Lactobacillus plantarum* strains. *Food Chem.* 107:320–326.
- Medina E, Brenes M, Romero C, Garcia A, de Castro A. 2007. Main antimicrobial compounds in table olives. *J. Agric. Food Chem.* 55:9817–9823.
- Barthelmebs L, Diviès C, Cavin J-F. 2000. Knockout of the *p*-coumarate decarboxylase gene from *Lactobacillus plantarum* reveals the existence of two other inducible enzymatic activities involved in phenolic acid metabolism. *Appl. Environ. Microbiol.* 66:3368–3375.
- Gury J, Seraut H, Tran NP, Barthelmebs L, Weidmann S, Gervais P, Cavin JF. 2009. Inactivation of PadR, the repressor of the phenolic acid stress response, by molecular interaction with Usp1, a universal stress protein from *Lactobacillus plantarum*, in *Escherichia coli*. *Appl. Environ. Microbiol.* 75:5273–5283.
- Gury J, Barthelmebs L, Tran NP, Diviès C, Cavin J-F. 2004. Cloning, deletion, and characterization of PadR, the transcriptional repressor of the phenolic acid decarboxylase-encoding *padA* gene of *Lactobacillus plantarum*. *Appl. Environ. Microbiol.* 70:2146–2153.
- Licandro-Seraut H, Gury J, Tran NP, Barthelmebs L, Cavin J-F. 2008. Kinetics and intensity of the expression of genes involved in the stress response tightly induced by phenolic acids in *Lactobacillus plantarum*. *J. Mol. Microbiol. Biotechnol.* 14:41–47.

25. Tran NP, Gury J, Dartois V, Nguyen TK, Seraut H, Barthelmebs L, Gervais P, Cavin JF. 2008. Phenolic acid-mediated regulation of the *padC* gene, encoding the phenolic acid decarboxylase of *Bacillus subtilis*. *J. Bacteriol.* 190:3213–3224.
26. Duy NV, Mader U, Tran NP, Cavin J-F, Tam le T, Albrecht D, Hecker M, Antelmann H. 2007. The proteome and transcriptome analysis of *Bacillus subtilis* in response to salicylic acid. *Proteomics* 7:698–710.
27. Mahillon J, Chandler M. 1998. Insertion sequences. *Microbiol. Mol. Biol. Rev.* 62:725–774.
28. Hayes F. 2003. Transposon-based strategies for microbial functional genomics and proteomics. *Annu. Rev. Genet.* 37:3–29.
29. Licandro-Seraut H, Brinster S, van de Guchte M, Scornec H, Maguin E, Sansonetti P, Cavin JF, Serron P. 2012. Development of an efficient *in vivo* system ( $P_{\text{junc}}$ -TpaseI $\Delta$ 223) for random transposon mutagenesis of *Lactobacillus casei*. *Appl. Environ. Microbiol.* 78:5417–5423.
30. Torriani S, Felis GE, Dellaglio F. 2001. Differentiation of *Lactobacillus plantarum*, *L. pentosus*, and *L. paraplantarum* by *recA* gene sequence analysis and multiplex PCR assay with *recA* gene-derived primers. *Appl. Environ. Microbiol.* 67:3450–3454.
31. Servili M, Baldioli M, Selvaggini R, Miniati E, Macchioni A, Montedoro G. 1999. High-performance liquid chromatography evaluation of phenols in olive fruit, virgin olive oil, vegetation waters, and pomace and 1D- and 2D-nuclear magnetic resonance characterization. *J. Am. Oil Chem. Soc.* 76:873–882.
32. Servili M, Urbani S, Esposto S, Taticchi A, Veneziani G, Maio I, Sordini B, Selvaggini R, Corsetti A. 2012. Standardizzazione del processo di deamarizzazione biologica per il miglioramento della qualità delle olive da tavola di varietà italiane, p 109–121. In Ed U (ed), Pogettazione, realizzazione e gestione di sistemi di controllo del rispetto delle norme di autenticità, qualità e commercializzazione dell'olio di oliva e delle olive da mensa immessi sul mercato. Tipografia Grafica G&G srl, Marino, Rome, Italy. (In Italian.)
33. Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
34. Aukrust TW, Brurberg MB, Nes IF. 1995. Transformation of *Lactobacillus* by electroporation. *Methods Mol. Biol.* 47:201–208.
35. Dower WJ, Miller JF, Ragsdale CW. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* 16:6127–6145.
36. Heap JT, Pennington OJ, Cartman ST, Minton NP. 2009. A modular system for *Clostridium* shuttle plasmids. *J. Microbiol. Methods* 78:79–85.
37. Duran MC, García P, Brenes M, Garrido A. 1993. *Lactobacillus plantarum* survival during the 1st days of ripe olive brining. *Syst. Appl. Microbiol.* 16:153–158.
38. Ruiz-Barba JL, Rios-Sanchez RM, Fedriani-Iriso C, Olias JM, Rios JL, Jimenez-Diaz R. 1990. Bactericidal effect of phenolic compounds from green olives on *Lactobacillus plantarum*. *Syst. Appl. Microbiol.* 13:199–205.
39. Juven B, Henis Y, Jacoby B. 1972. Studies on the mechanism of the antimicrobial action of oleuropein. *J. Appl. Bacteriol.* 35:559–567.
40. Newbury SF, Smith NH, Robinson EC, Hiles ID, Higgins CF. 1987. Stabilization of translationally active mRNA by prokaryotic REP sequences. *Cell* 48:297–310.
41. Reference deleted.
42. Wood JM. 2011. Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. *Annu. Rev. Microbiol.* 65:215–238.
43. Kramer R. 2010. Bacterial stimulus perception and signal transduction: response to osmotic stress. *Chem. Rec.* 10:217–229.
44. Tsakalidou E, Papadimitriou K. 2011. *Stress responses of lactic acid bacteria*. Springer, New York, NY.
45. Molenaar D, Bringel F, Schuren FH, de Vos WM, Siezen RJ, Kleerebezem M. 2005. Exploring *Lactobacillus plantarum* genome diversity by using microarrays. *J. Bacteriol.* 187:6119–6127.
46. Curiel JA, Rodriguez H, de Las Rivas B, Anglade P, Baraige F, Zagorec M, Champomier-Verges M, Munoz R, Lopez de Felipe F. 2011. Response of a *Lactobacillus plantarum* human isolate to tannic acid challenge assessed by proteomic analyses. *Mol. Nutr. Food Res.* 55:1454–1465.
47. Koponen J, Laakso K, Koskenniemi K, Kankainen M, Savijoki K, Nyman TA, de Vos WM, Tynkynen S, Kalkkinen N, Varmanen P. 2012. Effect of acid stress on protein expression and phosphorylation in *Lactobacillus rhamnosus* GG. *J. Proteomics* 75:1357–1374.
48. Di Cagno R, De Angelis M, Limitone A, Fox PF, Gobbetti M. 2006. Response of *Lactobacillus helveticus* PR4 to heat stress during propagation in cheese whey with a gradient of decreasing temperatures. *Appl. Environ. Microbiol.* 72:4503–4514.
49. Lee K, Lee HG, Pi K, Choi YJ. 2008. The effect of low pH on protein expression by the probiotic bacterium *Lactobacillus reuteri*. *Proteomics* 8:1624–1630.
50. Castaldo C, Vastano V, Siciliano RA, Candela M, Vici M, Muscariello L, Marasco R, Sacco M. 2009. Surface displaced alpha-enolase of *Lactobacillus plantarum* is a fibronectin binding protein. *Microb. Cell. Fact.* 8:14.
51. Rozes N, Peres C. 1998. Effects of phenolic compounds on the growth and the fatty acid composition of *Lactobacillus plantarum*. *Appl. Microbiol. Biotechnol.* 49:108–111.
52. van Bokhorst-van de Veen H, Abec T, Tempelaars M, Bron PA, Kleerebezem M, Marco ML. 2011. Short- and long-term adaptation to ethanol stress and its cross-protective consequences in *Lactobacillus plantarum*. *Appl. Environ. Microbiol.* 77:5247–5256.
53. Schulze-Gahmen U, Pelaschier J, Yokota H, Kim R, Kim SH. 2003. Crystal structure of a hypothetical protein, TM841 of *Thermotoga maritima*, reveals its function as a fatty acid-binding protein. *Proteins* 50:526–530.
54. Gibson TJ. 1984. *Studies on the Epstein-Barr virus genome*. Ph.D. thesis. Cambridge University, Cambridge, England.
55. Rochat T, Bouloc P, Repoila F. Gene expression control by selective RNA processing and stabilization in bacteria. *FEMS Microbiol. Lett.*, in press. doi:10.1111/1574-6968.12162.
56. Maldonado-Barragan A, Caballero-Guerrero B, Lucena-Padros H, Ruiz-Barba JL. 2011. Genome sequence of *Lactobacillus pentosus* IG1, a strain isolated from Spanish-style green olive fermentations. *J. Bacteriol.* 193:5605.

1 **Ability of *Lactobacillus pentosus* C11 and the "olive brine adaptation" derivated mutants to**  
 2 **form biofilms on olive skin during fermentation.**

3  
 4 Perpetuini G.<sup>1,2</sup>, Pham-Huang B. N.<sup>2</sup>, Scornec H.<sup>2</sup>, Tofalo R.<sup>1</sup>, Schirone M.<sup>1</sup>, Suzzi G.<sup>1</sup>, Corsetti A.<sup>1</sup>,  
 5 Waché Y.<sup>2</sup>, Cavin J.F.<sup>2</sup>, Licandro-Seraut H.<sup>2,3</sup>.

6  
 7 <sup>1</sup>Dipartimento Scienze degli Alimenti, Università di Teramo, Italy;

8 <sup>2</sup>UMR A 02102 PAM, AgroSup Dijon and Université de Bourgogne, Dijon, France;

9 <sup>3</sup>Unité de Pathogénie Microbienne Moléculaire, Institut Pasteur, Paris, France;

10  
 11 **Abstract**

12 *Lactobacillus pentosus* C11 is one of the few LAB strains capable of surviving in olive brine and  
 13 thus desirable during table olive fermentation. In this study, the ability to adhere and form biofilm  
 14 on olive skin was investigated for *L. pentosus* C11 and five derivative mutants unable to survive to  
 15 olive brine conditions. These five mutants are interrupted in metabolic genes (*enoA1* and *gpi*), and  
 16 in genes of unknown function (*oba* genes). Confocal microscopy together with live bacteria count  
 17 revealed that the sessile state represented the prevailing *L. pentosus* C11 life-style during table olive  
 18 fermentation. The characterization of cell surface properties showed that mutants presented less  
 19 hydrophobic and basic properties than the wild type (WT). In fact, their ability to adhere to both  
 20 abiotic (polystyrene plates) and biotic (olive skin) surfaces was significantly lower than that of the  
 21 WT. Confocal microscopy revealed that mutants adhered sparsely to the olive skin instead of  
 22 building a thin, multilayer biofilm. Moreover, qRT-PCR showed that the three genes *enoA1*, *gpi* and  
 23 *obaC* were upregulated in the olive biofilm compared to the planktonic state. Thus *enoA1*, *gpi* and  
 24 *oba* genes are necessary in *L. pentosus* to form an organized biofilm at the olive skin, the only state  
 25 allowing the bacteria to undergo a complete fermentation of olives.

26  
 27 **Introduction**

28 Olive fermentation is still carried out following artisanal practices despite its economic impacts. The  
 29 most significant production methods of table olives are: the Spanish-type and the Californian style  
 30 involving both a de-bittering process with sodium hydroxide solutions (Gómez et al. 2006), and the  
 31 Greek style where olives are directly brined and the debittering is slow and only partial (Tofalo et  
 32 al. 2012). Olive microbiota is mainly composed of lactic acid bacteria (LAB), yeast and  
 33 *Enterobacteriaceae* (Corsetti et al. 2012). Recently, a systematic use of suitable starter cultures,  
 34 mainly belonging to *Lactobacillus plantarum* and *Lactobacillus pentosus* species, was proposed to  
 35 initiate and to accelerate the fermentation process in order to prevent the development of  
 36 undesirable spoiling bacteria, and, *in fine*, to preserve the sensory quality of the final product  
 37 (Hurtado 2010; Arroyo-Lopez et al. 2012; Corsetti et al. 2012). In fact, these two species are able to  
 38 adapt, grow and survive in olive brine until the end of fermentation despite a low availability of  
 39 nutrients, a high salt concentration, a low pH, and the presence of phenolic compounds with  
 40 antimicrobial activity, which generate a global hostile and stressful environment.

41 To identify genetic mechanisms involved in the capacity of these species to face the olive  
42 fermentation conditions, we have previously carried out a global reverse genetic strategy on *L.*  
43 *pentosus* C11, a strain isolated from naturally fermented Itrana Bianca table olives (Perpetuini et al.  
44 2013). In particular, a library of 6000 random transposon mutants was generated and screened for  
45 the capacity to grow in a medium named BSM (Brine Screening Medium), displaying the stressful  
46 physical and chemical conditions encountered in olive brine. Five transposon mutants that had lost  
47 their ability to adapt and to grow in olive brine conditions were identified (Perpetuini et al. 2013).  
48 Genes involved in this adaptation are the metabolic genes *enoA1* and *gpi*, encoding respectively a  
49 putative enolase and a glucose-6-P-isomerase, and three genes of unknown function, which were  
50 named *oba* for “Olive Brine Adaptation”.

51 One of the main bacterial strategies to survive, adapt and even colonize new environment is surface  
52 adhesion and biofilm formation (Hall-Stoodley et al. 2004). A biofilm is a well organized,  
53 cooperating consortium of microorganisms, formed by one or more species, embedded in a matrix  
54 of extracellular polymeric substances (EPS) they have produced, and attached to either an inert or  
55 living surface (Wimpenny 2009; Remis et al. 2010). The benefits of growing in a biofilm  
56 environment include an increased ability to retain extracellular enzymes near the cell, protection  
57 from toxic compounds, increased resistance to stress conditions (Costerton et al. 1995; Lapaglia and  
58 Hartzell 1997). A first microscopic observation of bacteria and yeast adhering to black olive skin  
59 (Greek style) has been reported by Nychas et al. (2002). Thereafter, dynamics of a LAB population  
60 on olive skin (De Bellis et al. 2010) and a mixed-yeast/LAB population (Domínguez-Manzano et al.  
61 2012), were characterized. The *in situ* observation of olive skin showed that bacterial cells are  
62 embedded in an exopolysaccharide matrix, forming biofilms much more complex and consistent  
63 than on abiotic surfaces. Thus, this structure could enhance resistance to olive brine stresses,  
64 offering a physical barrier.

65 The aim of the present study was to assess the importance of biofilm formation ability for *L.*  
66 *pentosus* to complete table olive fermentation and to investigate a possible role of *oba* genes of *L.*  
67 *pentosus* C11 in the adhesion and subsequent biofilm formation on olive skin.

68

## 69 **Materials and methods**

70

### 71 **Bacterial strains and culture media**

72 *L. pentosus* C11 and the five mutants impaired for growth in olive brine (Tab. 1) were routinely  
73 grown at 37°C in MRS medium (Oxoid). For some experiments, these strains were cultivated in YG  
74 medium (10 g/L yeast extract, 10 g/L glucose) at pH 6. Strains were stored at -80°C in MRS broth  
75 supplemented with glycerol (20% v/v).

76

### 77 **Olive samples and brine**

78 Table olives, Itrana Bianca cultivar, were supplied by an artisanal producer located in Lazio region  
79 (Italy). The brine was taken from the end of fermentation of Itrana Bianca table olives (naturally  
80 fermented green table olives). This brine containing 70 g/L NaCl, and at pH 4.0, was mainly  
81 characterized by the presence of 3,4-DHPEA, p-HPEA, verbascoside, glucose and lactic acid  
82 (Servili et al. 2012; Tofalo et al. 2012). The brine-derived medium, used to grow bacteria in the

83 stressful conditions encountered in brine, was supplemented with 20 g/L glucose and 20 g/L yeast  
84 extract (Oxoid), adjusted to pH 6.0, and pasteurized at 65°C for 45 min.

85

#### 86 **Olive fermentations process and subsequent bacterial count**

87 *L. pentosus* C11 was subcultured overnight at 30°C in MRS broth containing 40 g/L NaCl for  
88 adaptation to the saline environment of olive brine. Bacteria were harvested by centrifugation  
89 (10000 × g), washed with PBS, and finally inoculated in three sterile glass jars, containing 300 mL  
90 pasteurized brine (70 g/L NaCl, pH 4) and 300 g of olives to a final concentration of 10<sup>7</sup> CFU/mL.  
91 Jars were maintained at 25°C for 2 months. The planktonic (non-attached) cells were incubated 1 h  
92 at 4°C, collected by centrifugation at 10000 × g, for 10 min at 4 °C and resuspended in 10 ml PBS  
93 (10 mM, pH 7.0). To collect sessile (attached) bacteria, olives samples (300 g) were washed three  
94 times in PBS and then placed in one liter erlenmeyer flask containing 300 mL of PBS and 10 g of  
95 glass beads (2 mm) (Biosigma) which was agitated on a shaker at 200 rpm for 1 h at 4 °C to detach  
96 microorganisms from olive skin. Then the liberated bacteria were harvested by centrifugation  
97 (10000 × g, 10 min, 4 °C) washed twice in PBS and suspended into 10 mL of PBS. Bacteria count  
98 were made on the harvest sessile and planktonic cells by plating serial dilutions on MRS. Also, the  
99 first washing from olive sample was plated and added to the count of planktonic cells, knowing that  
100 the bacterial count on the second and the third washing were negligible. The experiments were  
101 performed twice. The absence of contaminant on olive surface and in brine was checked by colony  
102 morphology and by multiplex PCR on the *recA* gene on randomly picked colonies, as previously  
103 described (Torriani et al. 2001).

104

#### 105 **Biofilm assay on polystyrene surface**

106 *L. pentosus* C11 and its mutants were grown in MRS at 37°C for 24 h. Bacteria cells were harvested  
107 by centrifugation at 5000 × g, and suspended in saline to yield an initial concentration of 10<sup>7</sup>  
108 CFU/mL. The ability to form stable biofilms was monitored by growing the cells in flat-bottom 6  
109 wells cell culture plates (Costar, Corning, NY) containing 5 mL YG medium at pH 6 for 72 h. Wells  
110 containing uninoculated growth medium were used as negative controls. Media and unattached cells  
111 were removed by rinsing with saline solution. The remaining bacteria were carefully rinsed with 1  
112 mL of saline and dispersed by pipetting up and down ten times as previously described (Kubota et  
113 al. 2009). Finally, the cell suspension was serially diluted in saline, and plated on solid MRS.  
114 Bacteria count obtained for 1 ml correspond to bacteria adhesion for one well. Three biological  
115 repeats were made.

116

#### 117 **Microbial adhesion to solvents (MATS) test**

118 The physico-chemical characterization of bacterial surface properties was evaluated by the MATS  
119 method (Bellon-Fontaine 1996) with some modifications. Bacteria in early stationary phase were  
120 harvested by centrifugation at 10000 × g for 10 min, washed twice and suspended in PBS to an  
121 OD<sub>600 nm</sub> of 0.8 (A0). A volume of 1 ml of solvent was added to 3 ml of cell suspension, and after a  
122 10 min-incubation at room temperature, the two phase system was mixed by vortexing for 5 min.  
123 After a 15 min-incubation at room temperature the aqueous phase was removed and its OD<sub>600 nm</sub> was  
124 measured (A1). The percentage of microbial adhesion to solvent was calculated as follow:  
125  $(1-A1/A0) \times 100$ . This procedure was applied to three different polarity solvents: hexadecane,

126 chloroform and ethyl acetate. The analysis of an eventual bacterial lysis under the action of these  
127 solvents was previously monitored by phase-contrast microscopy.

128  
129

### 130 **Colonization of olive skin, detachment of bacteria from olive skin and plate counts**

131 To test the ability of colonization of olive skin, olives were washed in 10 mM PBS to eliminate  
132 planktonic cells of the surface. Then, slices of olive skin (about 1 cm<sup>2</sup>) were meticulously peeled  
133 using a scalpel and each slice was added in a tube of 1 ml YG medium or brine-derived medium.  
134 *L. pentosus* was grown overnight in MRS broth, harvested by centrifugation, washed twice in PBS,  
135 suspended in PBS at 10<sup>7</sup> CFU/mL, and used to inoculate brine-derived medium or YG tubes  
136 cultures containing olive skin. Cultures were incubated at 25 °C for 10 days. Then, the olive skin  
137 was removed from tubes, washed three times in PBS, placed in tubes containing 0.5 g glass beads (2  
138 mm) and then vortexed for 5 min. Then, glass beads were eliminated, bacteria cells were collected  
139 by centrifugation and suspended into 1 mL of PBS. Samples of detached cells were serially diluted  
140 in saline solution (0.85% v/v NaCl) and plated for bacteria count. Three independent cultures were  
141 made for each condition. The absence of contaminant on olive skin was checked as described above.

142

### 143 **Confocal microscopy and image analysis**

144 Examinations of biofilms were performed using a confocal laser scanning microscopy Eclipse TE  
145 2000E (Nikon). The amount of attached bacteria was estimated by fluorescence microscopy after  
146 staining for 5 min with 2 µg/mL of 4',6-diamidino-2-phenylindole (DAPI D-9542, Sigma). Images  
147 were obtained with a Nikon Plan Apochromat VC 100x/1.4 oil-immersion numerical aperture (NA)  
148 objective. 3D image were recorded at an excitation wavelength of 408 nm by using an argon laser in  
149 combination with an emission long pass filter LP 505 nm. Digital image acquisition and analysis of  
150 the CLSM optical thin sections were performed with NIS Elements AR (version 4.10.01).

151

### 152 **RNA extraction and qRT-PCR analysis.**

153 After 2 months of olive fermentation with *L. pentosus*, planktonic and sessile bacteria were  
154 recovered (see "Olive fermentations process and subsequent bacterial count" for details). Total  
155 RNA was extracted using the Tri Reagent (Sigma) as previously described (Licandro-Seraut et al.  
156 2008). Quantitative Reverse Transcription-PCR (qRT-PCR) and calculations of relative transcript  
157 levels (RTLs) were carried out as previously described (Perpetuini et al. 2013), using the primers  
158 listed in Table 2, and the genes *tpiA* and *rpoD* as internal calibrators.

159

## 160 **Results**

161

### 162 **Balance between planktonic and sessile states of *L. pentosus* C11 cells during olive 163 fermentation.**

164 The localization of *L. pentosus* C11 during table olive fermentation and its ability to colonize olive  
165 skin were determined after inoculation of 3 x 10<sup>7</sup> CFU of *L. pentosus* C11 in a jar containing 300 ml  
166 of sterile brine and 300 g of washed olives of Itrana Bianca cultivar. Despite the fact that *L.*  
167 *pentosus* used here as the starter were pre-adapted to salt in MRS supplemented with 40 g/l NaCl,  
168 the population was decreased to 3 x 10<sup>5</sup> CFU/jar after 72 h of fermentation. After 2 months of

169 the population was decreased to  $3 \times 10^5$  CFU/jar after 72 h of fermentation. After 2 months of  
 170 fermentation, *L. pentosus* succeeded in survival and colonization in brine with values of  $6 \times 10^7$   
 171 CFU/jar of planktonic bacteria in the brine and  $2.4 \times 10^9$  CFU/jar of sessile bacteria on olive skin.  
 172 So, about 98% of *L. pentosus* cells adhered to olive skin, indicating that the sessile state represented  
 173 the prevailing *L. pentosus* C11 life-style during table olive fermentation. Confocal microscopy of 1  
 174  $\text{cm}^2$  slices of olive skin taken from the fermented olives obtained showed that a considerable  
 175 amount of *L. pentosus* C11 adhered to the olive skin and appeared embedded in a matrix which  
 176 keeps them in close proximity or attached to each other (Fig. 1), which is in agreement with  
 177 bacterial counts. Also, some bacterial cells appeared outside the matrix.

178

179 ***L. pentosus* C11 WT and “olive brine adaptation” mutants ability to form biofilm on abiotic**  
 180 **surface (polystyrene) and adhesion properties.**

181 Since sessile state was shown to be the main life-style during table olive fermentation, it was of  
 182 interest to compare the capacity of the wild and mutant strains to form biofilm on abiotic surface.  
 183 The strains were grown in crystal polystyrene 6-well plates containing YG medium during 72 h.  
 184 The mean cell numbers in biofilms of the WT was  $2.5 \times 10^8$  CFU/well. The three mutants  
 185  $\Delta\text{TTobaAB}$ ,  $\Delta\text{obaD}$  and  $\Delta\text{TTgpi}$  showed lower values of one Log compared to the WT (Fig. 2). For  
 186 mutants  $\Delta\text{TTobaC}$  and  $\Delta\text{enoA1}$ , the number of cells in biofilms was even lower (about 3 Log  
 187 lower), whereas growth rates of mutants were similar to WT for planktonic cells grown in the same  
 188 liquid medium. These results reveal a difficulty for all mutants to form biofilm on abiotic surface.

189 To evaluate the cell surface characteristics of WT *L. pentosus* and its mutants, MATS test was  
 190 applied (Fig. 2). WT *L. pentosus* display average affinities (about 50% adhesion) to chloroform (an  
 191 acidic, electron-donor solvent), ethyl acetate (a basic, electro-acceptor solvent) and hexadecane (a  
 192 hydrophobic solvent). All mutants present significant differences of surface properties, compared to  
 193 the WT. In more detail, all mutants, except  $\Delta\text{TTobaC}$ , display less basic (or electron acceptor)  
 194 properties than the WT. They are even less acid than basic since their affinity to ethylacetate is more  
 195 decreased than that to chloroform, compared to the WT. They are also less hydrophobic than the  
 196 WT.  $\Delta\text{TTobaC}$  has even lower affinities to all solvents than the other mutants.

197

198 ***L. pentosus* C11 WT and “olive brine adaptation” mutants ability to form biofilm on olive skin**

199 The 5 brine-sensitive mutants were assessed for their ability to form biofilm on olive skin after  
 200 incubation in either a rich medium (YG) or a brine-derived medium at pH 6, which contains the  
 201 brine chemicals (phenolics) and salinity but which permits the mutants to grow as the WT (Fig. 3).  
 202 Peeled olive skin samples (about  $1 \text{ cm}^2$ ) were covered with medium, inoculated with the WT or one  
 203 mutant, and incubated for 10 days. In brine-derived medium, the WT formed biofilm on the skin  
 204 with a cell density of  $5.1 \times 10^8$  CFU/ $\text{cm}^2$  after 10 days (Fig. 4). The mutants showed values ranging  
 205 from  $4.5 \times 10^3$  CFU/ $\text{cm}^2$  ( $\Delta\text{enoA1}$ ) to  $6.6 \times 10^4$  CFU/ $\text{cm}^2$  ( $\Delta\text{TTgpi}$ ), highlighting a difficulty in the  
 206 colonization process. The lower values were observed for  $\Delta\text{enoA1}$  and  $\Delta\text{TTobaC}$  mutants. Confocal  
 207 microscopy observation of olive skin in presence of olive brine revealed that, after 10 days, the WT  
 208 *L. pentosus* started to colonize olive skin forming a thin, multilayer biofilm. Also, some cells  
 209 appeared outside the matrix. Regarding mutants, they adhered sparsely to the olive skin. The same  
 210 approach was used to test ability of the WT and mutants to colonize olive skin in presence of YG  
 211 medium, so in absence of phenolic compounds characterizing olive brine and osmotic stress

212 triggered by salt concentration (70 g/l NaCl). WT and mutants adhesions in YG were similar to that  
213 obtained in brine-derived medium. These results suggest that *L. pentosus* colonize olive skin even  
214 in absence of stressful liquid environment.

215

#### 216 **Expression of “olive brine adaptation” genes in *L. pentosus* C11 sessile cells during olive** 217 **fermentation.**

218 To investigate the biofilm-related role of the interrupted genes, *L. pentosus* C11 was used as  
219 inoculum for olive fermentation of Itrana Bianca table olives at an initial cell concentration of about  
220  $10^7$  CFU/mL. After two months of fermentation, biofilms from the surface of the olives were  
221 detached using mechanical methods and qRT-PCR was performed, to compare olive brine  
222 adaptation genes expression in sessile and planktonic bacteria from the same fermented olive flask  
223 (Fig. 5). Genes *ftsh* and *luxS*, used as positive control since they are generally overexpressed in  
224 bacteria forming biofilms (Lebeer et al. 2007; Bove et al. 2012), were upregulated in sessile state  
225 comparing to the corresponding planktonic bacteria. Concerning olive brine adaptation genes,  
226 *enoA1*, *obaC* and *gpi* were upregulated in sessile bacteria by 33, 13 and 9-fold respectively.  
227 Expressions of *obaA*, *obaB*, *obaD* and *obaE* were not significantly changed.

228

#### 229 **Discussion**

230 Generally, bacterial attachment to surfaces depends on complex physicochemical interactions  
231 among the bacterial cells, the surface, and the liquid phase, which are caused by the cell surface  
232 charge, the hydrophobicity, and electron acceptor and donor properties (van Loosdrecht et al. 1987;  
233 Dickson et al. 1989; Briandet et al. 1999). The adhesion mechanism to olive skin is very complex  
234 and not completely understood due to the complexity of olive surface. Firstly, olive skin is coated  
235 by a highly-hydrophobic epicuticular wax. Secondly, it is the first place of olive release of nutrients  
236 but also of antimicrobials that should let microorganisms away. Some authors have observed that  
237 indigenous bacteria and yeast can form biofilm structures on olive skin (Nychas et al. 2002;  
238 Dominguez-Manzano et al. 2012). Epicuticular wax could provide a substrate on which biofilms can  
239 develop (Nychas et al. 2002) and thus resist to antimicrobials and benefit from the release of  
240 nutrients (Wimpenny 2009). By confocal microscopy analysis, we have revealed that olive surface  
241 provides a biotic support for *L. pentosus* C11 adhesion and development of living micro-  
242 architecture. This could explain the ability of this olive indigenous strain to face the olive brine  
243 stresses and complete table olive fermentations. In this study, we have also measured that sessile  
244 state represent almost the only life-style of *L. pentosus* C11 during olive fermentation (98% of the  
245 cells). The remaining planktonic cells could be considered as those that, detached from the biofilm,  
246 establish new communities and extend the biofilm, as commonly observed (Davey and O’Toole  
247 2000). Moreover, we have demonstrated that the presence of the stressful factors of the olive brine  
248 (NaCl and phenolic compounds) do not influence *L. pentosus* C11 ability to colonize olive skin at  
249 least during the first phases of fermentation since adhesion on olive skin is similar when the skin is  
250 incubated in YG medium and when it is incubated in brine-derived medium.

251 Characterization of the olive brine adaptation mutants allowed obtaining a deeper comprehension of  
252 the biofilm building at the olive surface. These five mutants presented modified surface properties,  
253 especially a decrease of hydrophobic and basic (electron-acceptor) properties. Therefore, their  
254 ability to adhere to polystyrene plates strongly decreased and, above all, they all lost their ability to

255 colonize olive skin. These two results could be related to the reduction of the hydrophobic character  
 256 of their cell surface since the plates used in this study are negatively charged and the olive  
 257 epicuticular wax is mainly made up of lipids comprising alkanes, alkyl esters, methyl phenyl esters,  
 258 aldehydes, triacylglycerols, alcohols, triterpenols, fatty acids and triterpene acids (Bianchi 2003)  
 259 thus generating repulsive electrostatic interactions between mutants and olive skin. Our results  
 260 indicates that all the *oba*, *enoA1* and *gpi* genes are required for biofilm formation on the olive skin,  
 261 at least because they contribute to maintain adequate cell surface properties, which is one of the  
 262 major determinant in the formation of biofilms on biotic and abiotic surfaces (Courtney et al. 2009).  
 263 Among these genes, qRT-PCR results indicate that *enoA1*, *obaC* and *gpi* were significantly  
 264 upregulated in the bacterial biofilm, compared to planktonic state, which suggest that they are  
 265 regulated to fit biofilm lifestyle. Genes *enoA1* and *gpi* encode two ubiquitous cytosolic enzymes  
 266 involved in the glycolysis pathway. Their up regulation could be explained by the fact that sessile  
 267 bacteria are more metabolically active than planktonic ones because they are closer to nutrients.  
 268 However, some studies revealed that their function can vary as a consequence of changes in their  
 269 cellular localization. In particular, the EnoA1 protein of *Lactobacillus plantarum* (a very close  
 270 species of *L. pentosus*) was localized at the cell surface and was capable of binding to human  
 271 fibronectin (Castaldo et al. 2009). Surface enolases presenting various binding specificity have also  
 272 been described in other LAB such as *Lactobacillus crispatus*, *Lactobacillus johnsonii* and  
 273 *Streptococcus suis* (Pancholi and Fischetti 1998; Pancholi 2001; Antikainen et al. 2007;  
 274 Hurmalainen et al. 2007; Esgleas et al. 2008). Gpi proteins have also been described as adhesive  
 275 moonlighting enzymes in *Lactobacillus crispatus* that associate on the cell surface in acidic  
 276 conditions (Kainulainen et al. 2012). Probably, Gpi and EnoA1 enzymes could be involved in the  
 277 adhesion to some specific components of olive skin. Regarding *obaC*, it encodes a putative fatty  
 278 acid binding protein of the DegV family which could bind lipids of the epicuticular wax.  
 279 This study proves that biofilm at the olive surface is the main lifestyle of *L. pensosus* C11 to achieve  
 280 olive fermentation. This olive-adapted strain is able to form biofilm despite the antimicrobial  
 281 compounds and the hydrophobic epicutular wax of the olives thanks to its surface properties and  
 282 probably with the help of surface proteins (ObaC, EnoA1, GpiA) that could bind olive surface  
 283 compounds. As *enoA1* and *gpiA* are also considered as genes of interest in beneficial host-bacterial  
 284 interaction, their up-regulation in sessile state on olive is one more claim to propose table olives as  
 285 functional food (Lavermicocca et al. 2005; De Bellis et al. 2010; Abriouel et al. 2011).

286

### 287 **Acknowledgement**

288 This work was supported by the scientific council of AgroSup Dijon. We are grateful to the Plateau  
 289 d'Imagerie Spectroscopique PIMS, DImaCell.

290

### 291 **References**

- 292 Abriouel H, Benomar N, Lucas R, Gálvez A (2011) Culture-independent study of the diversity of microbial  
 293 populations in brines during fermentation of naturally-fermented Aloreña green table olives. *Int J Food*  
 294 *Microbiol* 144:487–496.
- 295 Antikainen J, Kuparinen V, Lahteenmaki K, Korhonen TK (2007) Enolases from Gram-positive bacterial  
 296 pathogens and commensal lactobacilli share functional similarity in virulence - associated traits. *Fems*

- 297 Immunol Med Mic 51:526–534.
- 298 Arroyo-López FN, Bautista-Gallego J, Domínguez-Manzano J, Romero-Gil V, Rodríguez-Gómez F, García-  
299 García P, Garrido-Fernández A, Jiménez-Díaz R (2012) Formation of lactic acid bacteria–yeasts  
300 communities on the olive surface during Spanish-style Manzanilla fermentations. *Food Microbiol* 32:295–  
301 301.
- 302 Bellon-Fontaine NN, Rault J, van Oss CJ (1996) Microbial adhesion to solvents: a novel method to  
303 determine the electron-donor/electron-acceptor or Lewis acid-base properties of microbial cells. *Colloid*  
304 *Surface B* 7:47–53.
- 305 Bianchi G (2003) Lipids and phenols in table olives. *Eur J Lipid Sci Technol* 105:229–242.
- 306 Bove P, Capozzi V, Garofalo C, Rieu A, Spano G, Fiocco D (2012) Inactivation of the *ftsH* gene of  
307 *Lactobacillus plantarum* WCFS1: effects on growth, stress tolerance, cell surface properties and biofilm  
308 formation. *Microbiol Res* 167:187–193.
- 309 Briandet R, Meylheuc T, Maher C, Bellon-Fontaine MN (1999) *Listeria monocytogenes* Scott A: Cell  
310 surface charge, hydrophobicity, and electron donor and acceptor characteristics under different  
311 environmental growth conditions. *Appl Environ Microbiol* 65:5328–5333.
- 312 Cataldo C, Vastano V, Siciliano RA, Candela M, Vici M, Muscariello L, Marasco R, Sacco M (2009)  
313 Surface displaced alfa-enolase of *Lactobacillus plantarum* is a fibronectin binding protein. *Microb Cell Fact*  
314 8:14.
- 315 Corsetti A, Perpetuini G, Schirone M, Tofalo R, Suzzi G (2012) Application of starter cultures to table olive  
316 fermentation: an overview on the experimental studies. *Front Microbiol* 3:1–6.
- 317 Costerton JW, Lewandowski Z, Caldwell DE, Korber DR, Lappin-Scott HM (1995) Microbial biofilms.  
318 *Annu Rev Microbiol* 49:711–745.
- 319 Courtney HS, Ofek I, Penfound T, Nizet V, Pence MA, Kreikemeyer B, Podbielbski A, Hasty DL, Dale JB  
320 (2009) Relationship between Expression of the Family of M Proteins and Lipoteichoic Acid to  
321 Hydrophobicity and Biofilm Formation in *Streptococcus pyogenes*. *PLoS ONE* 4:e4166.
- 322 Davey ME, O'Toole GA (2000) Microbial biofilms: from ecology to molecular genetics. *Microbiol Mol Biol*  
323 *Rev* 64:847–867.
- 324 De Bellis P, Valerio F, Sisto A, Lonigro SL, Lavermicocca P (2010) Probiotic table olives: Microbial  
325 populations adhering on olive surface in fermentation sets inoculated with the probiotic strain *Lactobacillus*  
326 *paracasei* IMPC2.1 in an industrial plant. *Int J Food Microbiol* 140:6–13.
- 327 Dickson JS, Koohmaraie M (1989) Cell surface charge characteristics and their relationship to bacterial  
328 attachment to meat surfaces. *Appl Environ Microbiol* 55:832–836.
- 329 Domínguez-Manzano J, Olmo-Ruiz C, Bautista-Gallego J, Arroyo-López FN, Garrido-Fernández A,  
330 Jiménez-Díaz R (2012) Biofilm formation on abiotic and biotic surfaces during Spanish style green table  
331 olive fermentation. *Int J Food Microbiol* 157:230–238.
- 332 Gómez AHS, García PG, Navarro LR (2006) Elaboration of table olives. *Grasas y Aceites* 57:86–94.
- 333 Hall-Stoodley L, Costerton JW, Stoodley P (2004) Bacterial biofilms: from the natural environment to  
334 infectious diseases. *Nat Rev Microbiol* 2:95–108.
- 335 Hurmalainen V, Edelman S, Antikainen J, Baumann M, Lahteenmaki K, Korhonen TK (2007) Extracellular  
336 proteins of *Lactobacillus crispatus* enhance activation of human plasminogen. *Microbiol* 153:1112–1122.
- 337 Hurtado A, Reguant C, Bordons A, Rozès N (2010) Evaluation of a single and combined inoculation of a  
338 *Lactobacillus pentosus* starter for processing cv. Arbequina natural green olives. *Food Microbiol* 27:731–  
339 740.
- 340 Kainulainen V, Loimaranta V, Pekkala A, Edelman S, Antikainen J, Kylvaja R, Laakkonen L, Finne J,  
341 Korhonen TK (2012) Glutamine Synthetase and Glucose-6-Phosphate Isomerase are adhesive moonlighting  
342 proteins of *Lactobacillus crispatus* released by epithelial cathelicidin LL-37. *J Bacteriol* 194:2509-2519.

- 343 Kubota H, Senda S, Tokuda H, Uchiyama H, Nomura N (2009) Stress resistance of biofilm and planktonic  
344 *Lactobacillus plantarum* subsp. *plantarum* JCM 1149. Food Microbiol 26:592-597.
- 345 Lapaglia C, Hartzell PL (1997) Stress-induced production of biofilm in the hyperthermophile *Archaeoglobus*  
346 *fulgidus*. Appl Environ Microbiol 63:3158–3163
- 347 Lavermicocca P, Valerio F, Lonigro SL, De Angelis M, Morelli L, Callegari ML, Rizzello CG, Visconti A  
348 (2005) Study of adhesion and survival of lactobacilli and Bifidobacteria on table olives with the aim of  
349 formulating a new probiotic food. Appl Environ Microbiol 71:4233–4240.
- 350 Lebeer S, De Keersmaecker SCJ, Verhoeven TLA, Fadda A, Marchal K, Vanderleyden J (2007) Functional  
351 analysis of luxS in the probiotic strain *Lactobacillus rhamnosus* GG reveals a central metabolic role  
352 important for growth and biofilm formation. J Bacteriol 189:860–871.
- 353 Licandro-Seraut H, Gury J, Tran NP, Barthelmebs L, Cavin JF (2008) Kinetics and intensity of the  
354 expression of genes involved in the stress response tightly induced by phenolic acids in *Lactobacillus*  
355 *plantarum*. J Mol Microbiol Biotechnol 14:41–44.
- 356 Esgleas M, Li Y, Hancock MA, Harel J, Dubreuil JD, Gottschalk M (2008) Isolation and characterization of  
357 a-*enolase*, a novel fibronectin-binding protein from *Streptococcus suis*. Microbiol 154:2668–2679
- 358 Nychas GJE, Panagou EZ, Parker ML, Waldron KW, Tassou CC (2002) Microbial colonization of naturally  
359 black olives during fermentation and associated biochemical activities in the cover brine. Lett Appl  
360 Microbiol 34:173–177.
- 361 Pancholi V (2001) Multifunctional alpha-*enolase*: its role in diseases. Cell Mol Life Sci 58:902–920.
- 362 Pancholi V, Fischetti VA (1998) Alpha-*enolase*, a novel strong plasmin(ogen) binding protein on the surface  
363 of pathogenic streptococci. J Biol Chem 273:14503–14515.
- 364 Pelletier C, Bouley C, Cayuela C, Bouttier S, Bourlioux P, Bellon-Fontaine MN (1997) Cell surface  
365 characteristics of *Lactobacillus casei* subsp. *casei*, *Lactobacillus paracasei* subsp. *paracasei*, and  
366 *Lactobacillus rhamnosus* strains. Appl Environ Microbiol 63:1725–1731.
- 367 Perpetuini G, Scornec H, Tofalo R, Serror P, Schirone M, Suzzi G, Corsetti A, Cavin JF, Licandro-Seraut H  
368 (2013) Identification of critical genes for growth in olive brine by transposon mutagenesis of *Lactobacillus*  
369 *pentosus* C11. Appl Environ Microbiol 79:4568–4575.
- 370 Remis JP, Costerton JW, Auer M (2010) Biofilms: structures that may facilitate cell–cell interactions. ISME  
371 J 4:1085–1087.
- 372 Servili M, Urbani S, Esposto S, Taticchi A, Veneziani G, Maio I, Sordini B, Selvaggini R, Corsetti A (2012)  
373 Standardizzazione del processo di deamarizzazione biologica per il miglioramento della qualità delle olive  
374 da tavola di varietà italiane, p 109 –121. In: Pogettazione, realizzazione e gestione di sistemi di controllo del  
375 rispetto delle norme di autenticità, qualità e commercializzazione dell’olio di oliva e delle olive da mensa  
376 immessi sul mercato, Tipografia Grafica G&G srl, Marino, Rome, Italy. (In Italian)
- 377 Tofalo R, Schirone M, Perpetuini G, Angelozzi G, Suzzi G, Corsetti A (2012) Microbiological and chemical  
378 profiles of naturally fermented table olives and brines from different Italian cultivars. Antonie Leeuwenhoek  
379 102:121–131.
- 380 Torriani S, Felis GE, Dellaglio F (2001) Differentiation of *Lactobacillus plantarum*, *L. pentosus*, and *L.*  
381 *paraplantarum* by *recA* gene sequence analysis and multiplex PCR assay with *recA* gene-derived primers.  
382 Appl Environ Microbiol 67:3450–3454.
- 383 van Loosdrecht MCM, Lyklema J, Norde W, Schraa G, Zehnder AJB (1987) The role of bacterial cell wall  
384 hydrophobicity in adhesion. Appl Environ Microbiol 53:1893–1897.
- 385 Wimpenny J (2009) Microbial metropolis. Adv Microb Physiol 56:29–84.
- 386

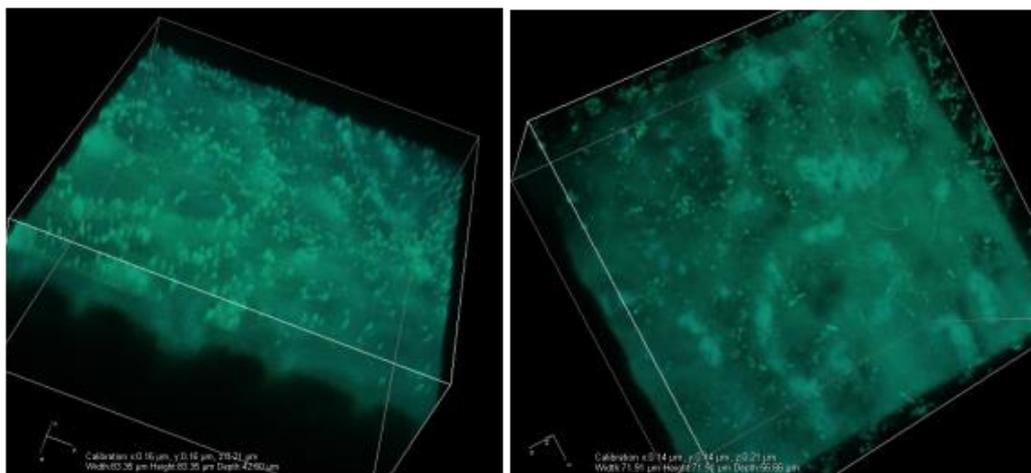


Fig.1. 3-D reconstructed images showing *L. pentosus* C11 biofilms formed on olive surface

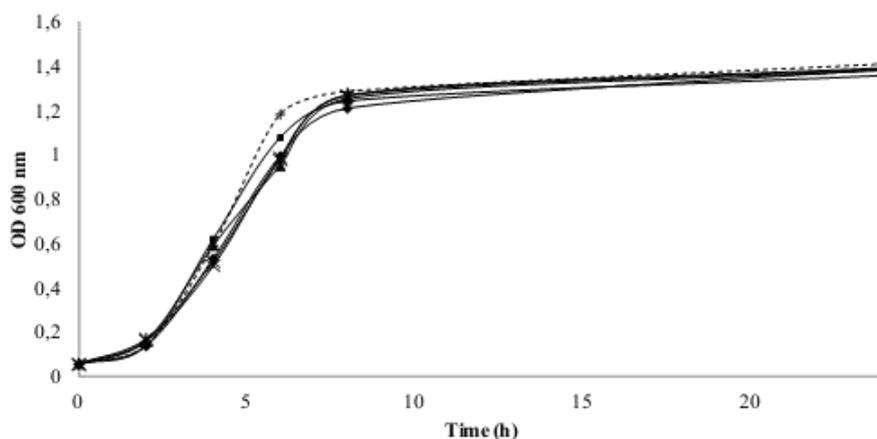


Fig. 3. Growth of WT *L. pentosus* C11 (- + - -) and brine sensitive mutants  $\Delta TTobaAB$  (■),  $\Delta obaD$  (▲),  $\Delta TTgpi$  (×),  $\Delta TTobaC$  (○),  $\Delta enoA1$  (□) in the modified brine

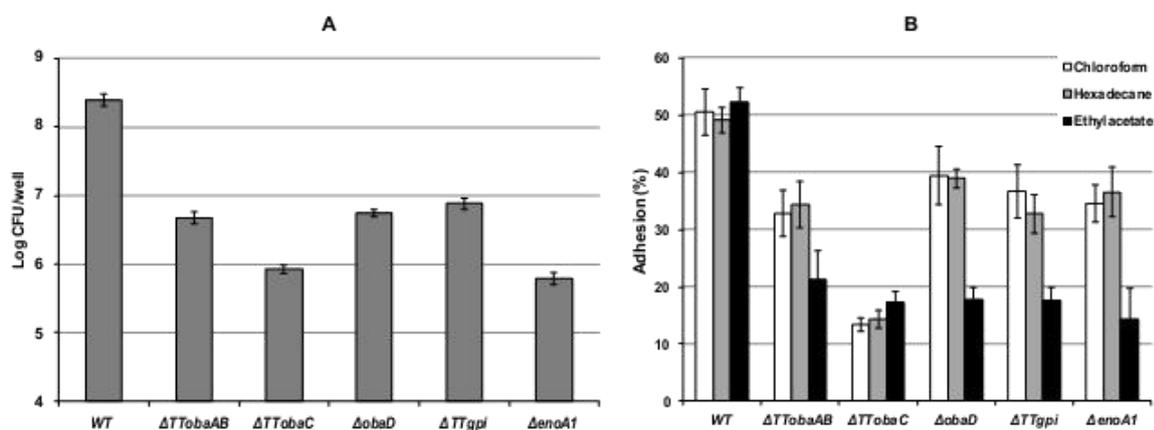
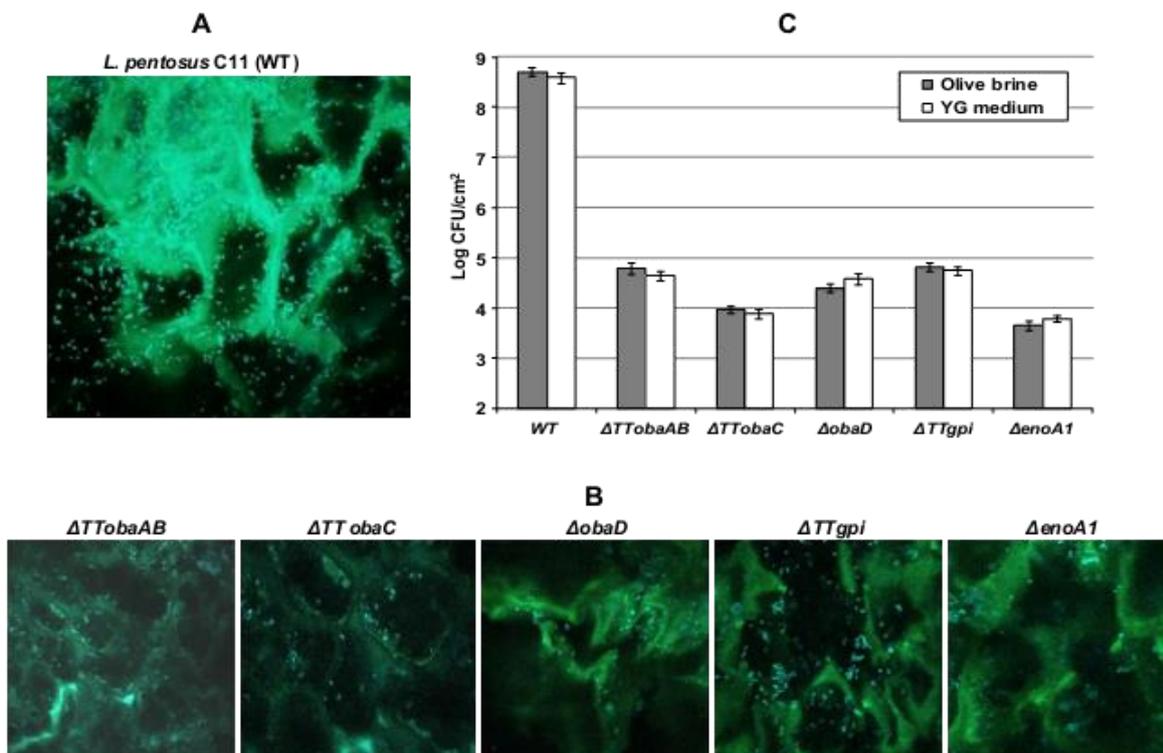
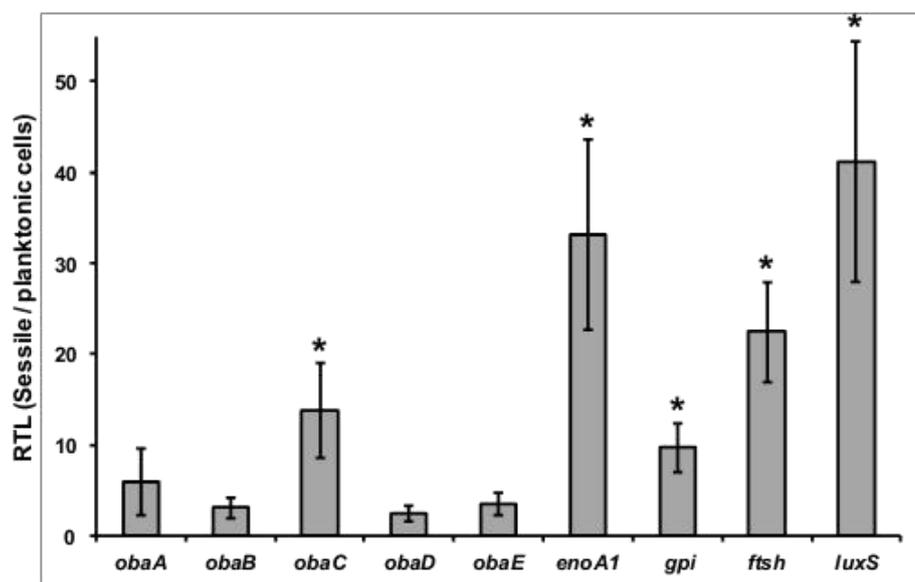


Fig. 2. A, Viable cell counts of bacteria aggregated in biofilms on polystyrene surface; B, Adhesion of *L. pentosus* C11 and its mutants to hexadecane, chloroform and ethylacetate



**Fig. 4.** Colonization of olive skin by *L. pentosus* C11 and mutants. **A, B.** Confocal microscopy observations of olive skin after incubation for 10 days in modified olive brine with *L. pentosus* C11 (A) or mutants (B). **C.** Enumeration of *L. pentosus* C11 and its mutants adhering to olive skin in presence of modified olive brine or YG medium (results obtained from biologic triplicate).



**Fig. 5.** Relative transcript levels of *L. pentosus* C11 genes after two months of fermentation with olive in brine. Transcript levels of each gene are expressed as the relative fold change, with planktonic cells as the reference condition (fold change=1). Four biological repeats were performed, and bars indicate standard deviations; \*,  $p < 0.05$